Ethic Guidelines
for Applying Recommender Technologies
and Artificial Intelligence-Based Algorithms in Digital Services

I. General

What Are These Guidelines for?

These Guidelines for Applying Recommender Technologies and Artificial Intelligence-Based (hereinafter, the AI) Algorithms in Digital Services (hereinafter, the Guidelines, the Document) have been developed in furtherance of the AI Code of Ethics and are designed to systematize ethical practices and summarize approaches that are permissible and recommended for use when developing and configuring recommender algorithms, as well as when deploying and operating the functionality of personalized recommendations in digital services of various types.

The Guidelines are suggested for voluntary implementation by companies using recommender systems in digital services.

Terms

The Guidelines define the key terms as follows:

- **Digital Service** means an information resource that provides users with remote access to information, goods or services.
- **Hypothesis** means a reasoned assumption about the causes and interrelations in respect of the user's preferences.
- **Personalized Recommendations** means functional elements of the service interface showing the user one or more recommendation objects based on mathematical probabilistic computations of the service recommender system of that user's preferences in order to improve his/her user experience and to enhance satisfaction with the service.
- **Recommendation Object** means information, materials, goods or other objects chosen by the recommender system algorithm.
- **Recommender System Algorithm** *(hereinafter also: recommender technologies and algorithms, recommendations, recommender technologies, recommender algorithms)* means an automated system of consistent operations to choose information or other service objects that may meet the preferences of a particular user based on the history of such user's interaction with the service.

.

- **Recommender System** means software or hardware integrated into the service, being an integral part thereof or a separate system that implements the functionality of personalized recommendations based on mathematical algorithms.
- **Transaction** means a user's resulting action with respect to the objects presented in the service (viewing, evaluating or commenting on a material or publication, ordering a service or purchasing goods, etc.).
- **User** means an individual using the digital service including the functionality of personal recommendations.
- **User Agreement** means an agreement between the user and the owner (administrator) of the service on the terms and conditions of its use that is usually entered into by the user accepting an offer.
- **User Data** means information on the users that is available, accumulated by the services or communicated directly to the services by the users.
- **User Interface** means organized and compiled display of information, decorative elements and elements controlling the service functionality.
- **User Preferences** means a set of data on information, materials, goods or other objects that the user interacted with while using the service.

Role of Recommender Systems in Digital Services

Modern digital services typically contain significant amounts of information or represent a wide range of goods and services that, taken together, meet individual needs of a large number of users. The service key task is to generate the most relevant response to the user's request, regardless of how accurately the request is worded. How successfully this task is solved determines directly the economic soundness and competitiveness of the service, as each successful user transaction is commercialized in one way or another.

As a rule, the materials, information, goods or services in the services are catalogued or otherwise structured to make it easier for the user to find what is needed. The context search functionality provides for finding the object searched not only by its name, but also by relevant keywords.

If the user was able to meet his/her need using the service, and that need is regular, there is a high probability that the user will access this service again or on a regular basis. While the user repeatedly accesses the service, the pattern of his/her preferences, i.e. range of his/her needs, sphere of interests, likings will be gradually shaped on the basis of features of the objects which he/she interacts with. The combination of user preferences data and data describing those users themselves makes it possible to generalize user groups' preferences.

.

Preference analysis makes it possible to make hypotheses about the user's needs at the moment of his/her next use of the service. Based on those hypotheses, the service forms and communicates recommendations, i.e. a set of objects that are highly probable to satisfy the user's needs, to the digital service user. If the hypotheses prove to be true and the user makes successful transactions with the proposed objects, i.e. satisfies his/her needs with minimal labor consumption (the appropriate object will not need to be searched for in the catalog or using the search function), this will form successful experience. The successful experience with the service not only contributes to the long-term commercialization of repeated accompanying transactions, but also builds the user's loyalty to the service, i.e. the practically confirmed confidence that this service is optimally suited to meet his/her particular needs. The stable flow of commercialized transactions and loyal audience are key factors in the economic efficiency of digital services, and the competition of services contributes to increasing consumer satisfaction, i.e. improving the quality of life in the society.

In order to concurrently make a lot of hypotheses about the individual needs of many users, artificial intelligence and mathematical algorithms are used on the basis whereof the services implement the functionality of personalized recommendations. Thus, the recommender systems are today an integral part of most digital services. Using recommender systems significantly reduces transaction costs for the user on the way to the target. Such optimization significantly improves the quality of life of the consumer and the economic efficiency of manufacturers and suppliers of information, goods and services.

Types of Recommender Systems

There are four basic types of recommender systems. The use of certain types in digital services of various areas is subject to technical capabilities of these services and what user data they have.

- *Content-Based*
  Such recommendations are based on the coincidence of features of the objects which the user successfully interacted with (goods trademark, brand or model of clothes, plot or genre of motion pictures, style or author of music, etc.). Such recommender systems are widely used in e-commerce services, digital content access services, social media and video hosting.

  E.g., modern marketplaces assign to the goods more than 100 features with various significance (weight) that are taken in account by the algorithms while generating recommendations.

.

- *Collaborative Filtering*
  This type is based on the hypothesis that users who are similar to each other may share similar preferences. It is most widely used in e-commerce services.

  > *E.g., the Worth a Look recommender block of Sbermarket contains goods that the user bought earlier or goods that were bought by other users similar to him/her.*

- *Knowledge-Based*
  This type functions on the basis of knowledge of a particular industry domain and the compatibility of objects, such as recommending spare or auxiliary components to sophisticated technical devices. It is also most widely used in e-commerce services.

  > *For example, this type of recommender services is actively used in dedicated online stores: Technopark, M.Video, Avito, etc.*
  > *So, Avito shows cards of accompanying (supplemental) goods at the bottom of the card of viewed goods:*
  >   - *The user is choosing a phone, and offers with cases for the phone model that the user is looking at are shown at the bottom of the phone card.*
  >   - *The user is choosing a coffee machine, and offers with filters for the coffee machine model that the user is looking at are shown at the bottom of the coffee machine card.*

- *Hybrid*
  Concurrent use of several types of recommender systems. Similar (content-based and collaborative model). This block helps the user to find goods similar to those in the current goods card as quickly as possible, e.g.:
  - the same product, but with a different taste
  - a similar product, but from another brand
  - a similar product of the same brand, but with a different volume.

II. Guidelines
for Ethics of Applying Recommender Technologies
and Artificial Intelligence-Based Algorithms in Digital Services

Transparency and Openness of Operation Principles

In digital services, it is recommended to publicly announce the fact of recommender system application and to disclose the objectives and the general principles of their operation. Such disclosure may be made in the user agreement or otherwise communicated to the user.

> *E.g., OZON disclosed the recommender system algorithms of the Search and Recommendation blocks on the website in 2022. Zen users now understand the process of recommender algorithm operation, and they clearly see why they are offered particular content in the feed.*

Reasonable restrictions on the disclosure of such information may be appropriate based on the specifics and particular features of the business model of a particular digital service or subject to statutory restrictions, e.g., those related to intellectual property protection.

> *Comment: it is not feasible to fully disclose the marketplace recommender system algorithms, otherwise, sellers will take advantage of vulnerabilities in the algorithms to artificially promote their goods.*

If such substantiated restriction exists, it is recommended that a provision on transparency and openness of the recommender system be implemented to the greatest extent possible.

It is also recommended to use the feedback functionality granting to the user an opportunity to leave a feedback or to ask a question about the operation of recommender systems.

> *E.g., many organizations have already implemented this function in the format of product feedback, contact us, wishes and comments and in the format of feedback on the processing of consumer personal data on the website.*

.

Age Restrictions

If the service is intended for or used, *inter alia*, by minors, and potential objects of the recommendations include any materials, goods and services that are not intended for or are prohibited for dissemination among minors, the recommender system of such service must abide by age restrictions when displaying information.

> *E.g., SberMarket does not show 18+ goods (alcohol available for ordering through self-delivery) in its recommender blocks. A similar approach is taken at other similar marketplaces.*

Users' Right to Revise Their Preferences

It is recommended that the services enable the user to revise his/her preferences, *inter alia*, by explicitly expressing no interest in certain objects ("not interested" button) or deleting certain objects from the history of his/her interaction therewith ("delete from the view history" button).

> *E.g., a SberMarket user can mark any goods as favorite ones and cancel such choice at any time. A Zen user can mark any information block as uninteresting with the "hide all publications of the channel" button, so this source of information will immediately disappear from the feed, and the priority of this information type will be reduced.*

Where technically feasible and reasonable subject to the intended purpose of the recommender system, it is recommended to grant to the users an opportunity to disable the recommender system in full or in part when using the digital service.

> *Comment: the full or partial disabling of the recommender system must be in line with the technical capabilities of the platform and the service purpose. A service specifically created solely to provide personalized content/goods/services should not implement this function, otherwise it would be contrary to the service idea and purpose.*

.

Use of User Data

It should be taken into account that the use of user data for the operation of recommender services should be based on the personal data protection laws and meet the requirements of other regulations.

> *E.g., each organization that provides any online services to customers, inter alia, using recommender algorithms, describes the procedure for handling user personal data on its website in the Privacy Policy, Policy for Collecting Personal Data on the Website sections as required by Federal Law No. 152-FZ dated 27 July 2006 On Personal Data.*

At that, subject to the intended purpose of the recommender system and legal restrictions, including those on the protection of rights to intellectual deliverables, it is recommended to disclose the types of user data used in the operation of recommender systems, either publicly or at the request of the user.

Restriction on Filter Bubbles

When using the recommender services, it should be taken into account that one of the key problems with their application is the phenomenon of so-called filter bubbles. This widely studied phenomenon is the consequence of people's desire to affirm their own opinions and positive self-esteem, which can form the so-called cognitive distortion. The recommender system algorithms can significantly increase the effect of such distortions. When the content-based recommender systems filter materials based on expressed user preferences and taking this feedback into account in a cycle, the diversity of features of filtered objects narrows gradually but steadily. The algorithm operation artificially narrows the user's range of vision, "preserves" his/her diet, prevents his/her development and blocks his/her obtaining new information, learning about new goods, services, etc.

Developers of recommender systems are advised to take a responsible approach to the filter bubble problem, study it, assess the scope of its occurrence and its consequences, and adjust the algorithms accordingly so as to mitigate the probability of its occurrence.

One way to solve this task is by using hybrid recommender systems that will support the diversity and relevance of recommendations.

.

Recommendations Evolving into Solicitation

When implementing content-based recommender systems, one should take into account the probability of a user having a repeated need for a particular object. An explicit mismatch between object consumption and recommendation cycles may be perceived by the user as solicitation and irritate him/her.

> *E.g., it is appropriate to recommend listening to a favorite piece of music once again during the day. At the same time, an offer to the user to buy once again an expensive non-deteriorating product during the day is controversial. To solve this problem, content-based recommender systems can be combined with knowledge-based recommendations.*

Target Audience and Sensitive Topics

It should be borne in mind that the use of recommender services to offer users goods and services that are explicitly designed for a particular target audience may be negatively perceived outside of that audience.

It should be taken into account that a number of goods and services may be highly sensitive, and the use of recommender services to promote such goods and services should be restricted. These are, e.g., goods and services:

- for adults (18+)
- of a religious nature
- of a ritual nature
- which may spark conflicts or ethnic hatred.

Recommendations Related to User Risk

Recommendations of goods and services involving risk should be treated in a special way. Although the current legislation as a rule establishes specific requirements for the sale of goods and provision of services involving risk, the visual presentation of such objects of recommendations should be accompanied with relevant warnings and the user agreement should set the limits of liability of the service and the user. Such recommendations may be relevant for services that provide recommendations on risky investment.

Anti-Fraud and Combating Manipulation of Algorithms

The system of relations among parties in the platform services is much more complex than that in conventional services, as the information, goods and service providers using such platform compete for user attention. It should be also noted that

.

the platform services usually ensure the interaction among a significant number of subjects in combinations of b2b, b2c and c2c.

In such situation, the service platform inevitably takes on the role of an arbitrator to maintain sound suppliers' competition and to combat abuse. In most cases, this approach is commercially justified by the business model of interaction between the service platform and suppliers.

In turn, unless the law provides for otherwise, the status of an information intermediary and the limited liability of the service platform for the quality and legality of information, materials, goods and services provided by a third party do not oblige the owner (administrator) of such service platform to warrant to the user the quality and legality of the provided and recommended objects.

The use of recommender systems and their significant impact on the mass demand draws to the recommender algorithms particular attention of manufacturers and suppliers that are extremely interested in getting their materials, goods and services into a sample of recommended objects for as many users as possible. If the recommender system algorithm is vulnerable to actions artificially simulating high interest of the platform users, this may result in an undesirable situation when objects are recommended with objective consumer properties (including value for money ratio) that are knowingly worse than those of other goods and services presented on the service platform. Such simulation of interest of a great number of users (fake engagement) can be achieved, among other things, automatically using specialized fraud systems and bots, user groups and even dedicated AI-based systems.

Distortion of information on the number of real users of the information resource can significantly affect the exercise of rights and obligations of third parties.

Currently, there is a trend to manipulate the results of measurement of viewing or listening to the content on video hosting and streaming platforms. Unfair persons intentionally make fake engagement of viewing/listening to certain works by various technical means, which affects the allocation of funds among sites and right holders. Fair right holders earn less payments, while unfair persons that have artificially increased their rating earn more.

Manipulation with digital objects, including large-scale simulation of natural activities and user profiles through automated processes and manipulative accounts threaten consumers' and companies' trust in digital goods and services.

In the course of such manipulation, audio and video content is played back by an individual or through machine means but it is not viewing or listening by real users. Streaming services can be manipulated in relation to individual works or groups of records in order to artificially improve positions on charts, increase a

.

market share, increase royalty contributions or for any other fraudulent or unfair purposes.

Service manipulation can not only cause economic harm to service providers, right holders, artists and advertisers, but also distort the mass media's and fans' impressions and their understanding of the popularity of individual content units, and harm the consumer experience and enjoyment of services by affecting the results of algorithmic reproduction. In addition to the economic harm, it harms right holders as well by providing them with potentially misleading and artificial data.

Thus, the manipulation of content viewing and listening distorts real user preferences.

Factors proving the manipulation include the following ones:
- an unusually high share of either free or premium views
- an unusually high share of certain types of devices
- an unusually high share of mobile devices as compared to desk-top ones
- a suspicious number of skips (in general and/or for individual customers)
- suspicious duration of playing back (e.g., a track is permanently skipped after 31 or 53 seconds, in general and/or for individual customers)
- an unusual number of play-backs by one individual user (a date/time/duration mark)
- all views are from one particular play-back list
- users whose behavior is suspicious in terms of location, time of day they broadcast (e.g., in the middle of the night, whereas they used to receive content during the day only).

Combating such fraud, fake engagement and any other forms of external manipulation of recommendations when it is not related to advertising is a major challenge for all platform services. Service owners are advised to take a comprehensive approach that includes the use of information security systems, monitoring and filtering of suspicious user actions taken from the same IP address or having signs of automatic generation, as well as other administrative measures aimed at discouraging counterparties and users to abuse. It is also advised to respond to messages of persons whose rights and legitimate interests may have been violated in the course of algorithm manipulation.

Taking User Feedback, Assessments and Complaints into Account

The recommender systems used in the social service platforms can have a material impact on the virus dissemination of information some of which may contain unreliable or even publicly dangerous information, cause anxiety or other damage. At that, such information may fail to meet the criteria that clearly classify it as prohibited.

.

Implementation of the complaints functionality may be one of the approaches to mitigate the risk of rapid dissemination of destructive information. Any users who have identified a destructive or dangerous nature of the material should be able to form a complaint or to mark it accordingly. In turn, recommender system algorithms should take into account these complaints, marks, and other forms of assessing the negative potential of particular materials when including them in recommendations, *inter alia*, block any objects that have received a critical number of user complaints from getting into recommendations, even despite the total number of signs of widespread user interest in such object.

User Health Recommendations

Where applicable, recommendations may be deployed that invite the user to take certain actions to keep up his/her health, including interrupting the service use for warm-up, rest or sleep.

> *E.g., suppose that a restaurant delivery aggregator notices that a consumer orders fatty fast food only. Is it ethical to expressly offer him/her a healthier food? Netflix has the "Are you still watching? functionality, the question appears when continuous viewing lasts for several hours. If no response is received within a few minutes, the play-back will stop as the service believes that the viewer has fallen asleep. The BMW onboard computer offers the driver to have a rest or coffee break after several hours of continuous driving.*

.

III. What Else Should Be Considered When Using Recommender Services

Legislative Regulation

The owners (administrators) of digital services focused on the users (consumers) from the Russian Federation should take into account that the use of recommender systems is governed, *inter alia*, by the following regulatory legal acts:

- Federal Law No. 408-FZ dated 31 July *2023 On Amendments to the Federal Law On Information, Information Technologies and Protection of Information*
- Law of the Russian Federation No. 2300-1 dated 07 February 1992 *On Consumer Rights Protection*
- Federal Law No. 149-FZ dated 27 July 2006 *On Information, Information Technologies and Protection of Information*
- Federal Law No. 381-FZ dated 28 December 2009 *On Fundamentals of State Regulation of Trade in the Russian Federation*
- Federal Law No. 152-FZ dated 27 July 2006 *On Personal Data*
- Federal Law No. 436-FZ dated 29 December 2010 *On Protection of Children from Information Causing Harm to Their Health and Development*
- Law of the Russian Federation No. 2124-1 dated 27 December 1991 *On Mass Media*
- Federal Law No. 38-FZ dated 13 March 2006 *On Advertising* (as amended)
- Federal Law No. 15-FZ dated 23 February 2013 *On Protection of Public Health from Exposure to Ambient Tobacco Smoke, Consequences of Tobacco Consumption or Consumption of Nicotine-Containing Products*
- Federal Law No. 135-FZ dated 26 July 2006 *On Protection of Competition*, etc.

The list of regulatory legal acts above is not exhaustive. The persons applying the recommender systems must be aware of their liability and comply with the rules of law.

The owners (administrators) of the services focused on the users (consumers) located, *inter alia*, outside the Russian Federation, should correlate their activities with the national legislation.

Direct and Indirect Recommendations

Various digital services are characterized by both general and specific issues of ethical use of the recommender systems based on the type of digital service, which can be divided as follows:

.

- **The service** disseminates materials and information or provides goods and services in the first person. At that, the service owner (administrator) is fully liable for all materials, information disseminated and for the legality and quality of goods and services sold. The most accurate comparison is the mass media with editorial offices producing and being liable for the materials published.
- **The service platform** provides its digital platform to third parties for the provision of goods and services, exchange of information, and dissemination of information and materials. At that, the liability for the material and information disseminated through such platform and for the legality and quality of the goods and services sold is borne by their authors and suppliers, while the platform owner (administrator) is not liable or bears limited liability according to the information intermediary status or liability established by law separately.

Restrictions on Application of These Guidelines

The recommender algorithms used in the service are applied to objects represented in that service only. In this regard, ethical and other issues contained in the information and materials that could be the object of recommendations should be distinguished from any issues that could result from deficiencies in the recommender system algorithms.

These Guidelines may not apply to practices and relations in the advertising, including personalized advertising.

These Guidelines are not applicable to methods and approaches to editorial or administrative (not automated) generation of recommendations.

Any practices and approaches not included in the Document, as well as any reasonable deviations from the provisions of the Document, may not be recognized to be contrary to the principles of ethics, prudence, good faith, and fairness for that reason only that such practices are not included in the Guidelines or do not fully comply with the provisions of the Guidelines.

The Document does not provide any ethical or legal assessment of information, materials, goods and services that may become the object of recommendations on the basis that the requirements for dissemination of information, turnover of certain goods, and provision of certain services, *inter alia*, through digital services, are governed by relevant laws, and the timely identification and suppression of violations of law is the competence of authorized public bodies.

.