

УДК 004.652

DOI: [10.26102/2310-6018/2023.42.3.007](https://doi.org/10.26102/2310-6018/2023.42.3.007)

Стандарт передачи характеристик товаров

А.Г. Шишкин✉

Группа компаний “Икс-Ком”, Москва, Российская Федерация

Резюме. Бизнес показатели любого интернет-магазина напрямую зависят от качества товарного контента. Получение такого контента часто пытаются автоматизировать, но на рынке пока не существует единого формата обмена такими данными. В итоге компании тратят большие ресурсы на интеграции с разными поставщиками и потребителями контента. Автор статьи предлагает формат для стандартизации процесса обмена данными с учетом особенности работы всех типов участников рынка электронной коммерции. Формат основан на опыте разработки технологической платформы для крупного интернет-магазина и его интеграции с партнерами. Учитываются особенности разных сегментов рынка, типов товаров, исторических данных, существующих форматов обмена, международных стандартов, тренда развития рынка. Формат разработан так, что любой представитель рынка может начинать его применять, не проводя серьезных доработок внутри имеющихся систем, используя модули-конвертеры, базируемые на чтении нужной глубины детализации товарного контента. Подобная стандартизация позволит сокращать издержки всех участников рынка при обоюдной интеграции, быстро распространять качественный контент, верифицировать его, сокращать число ошибок. Это может привести к росту конкуренции, так как малые участники рынка смогут получать равный доступ к данным. Все это в конечном итоге позволит предложить покупателям более качественный сервис и снижение цен на товары и услуги.

Ключевые слова: товарный контент, база данных, аналитика контента, контроль ошибок контента, управление качеством, карточка товара, электронный паспорт товара, электронная коммерция, передача контента.

Для цитирования: Шишкин А.Г. Стандарт передачи характеристик товаров. *Моделирование, оптимизация и информационные технологии.* 2023;11(3). URL: <https://moitvivr.ru/ru/journal/pdf?id=1388> DOI: 10.26102/2310-6018/2023.42.3.007

Product specification transfer standard

A.G. Shishkin✉

X-Com Group of Companies, Moscow, the Russian Federation

Abstract. The economic indicators of any online store directly depend on the quality of product content. The acquisition of such content is often automated, but there is no single data exchange format on the market yet. As a result, companies spend a lot of resources integrating with various content suppliers and consumers. The author of the article proposes a format for process standardization taking into consideration the features of all types of e-commerce market participants. The format is based on the experience of developing a technological platform for a large online store and its integration with partners. The features of various market segments, types of goods, statistical data, existing exchange formats, international standards, market development trends are accounted for. The format is designed so that any market representative can start using it by means of the converter modules based on reading the required speed of commodity consumption detail without major modifications inside the access system. Such standardization will help to cut costs for market participants while integrating mutually, quickly distribute quality content, verify it, reduce the number of errors. This may lead to increased competition as small market participants will have equal access to data. Ultimately, all of this will enable vendors offer their clients higher quality services and reduced prices.

Keywords: product content, database, content analytics, content error control, quality management, product card, electronic product passport, e-commerce, content transfer.

For citation: Shishkin A.G. Product specification transfer standard. *Modeling, Optimization and Information Technology*. 2023;11(3). URL: <https://moitvvt.ru/ru/journal/pdf?id=1388> DOI: 10.26102/2310-6018/2023.42.3.007 (In Russ.).

Введение

Рынок электронной коммерции имеет огромный потенциал роста как в РФ, так и в мире [1]. Вся электронная коммерция базируется на информации о товарах. Общая мировая товарная матрица растет с каждым днем, развиваются технологии и технологичность товаров, появляются крупные и очень крупные интернет-ресурсы по агрегации товарных матриц [2-4]. Вместе с этим растет и конкуренция среди интернет-магазинов, которым важно иметь качественный контент на их сайтах [5]. Все это требует и развития новых, более сложных систем учета и хранения товаров, в том числе детального учета характеристик товаров [6]. Говоря иными словами, требуется стандартизация.

На мировом рынке уже есть попытки сделать такие системы [7, 8]. Например, есть проект IceCar.biz, который заявляет наличие у него 10 000 000 карточек товара. На практике же оказалось, что у него набор характеристик по товарам крайне скудный, а вот формат хранения характеристик уже представляет некоторый интерес.

Широкое применение получили PIM (Product information management system, системы управления товарным контентом) – системы как от российских разработчиков, так и от международных [9]. Однако изучение их возможностей, модели хранения данных, моделей экспорта / импорта, показывает, что все они работают в разных форматах. Каждая система сама решает, как именно они будут хранить свои данные. Это в конечном счете вынуждает потребителей проводить сложные интеграции в свои системы. Стоит отметить, что сама по себе интеграция не так сложна, как сложен процесс мапинга (синхронизации) всех справочников, полей и значений.

Похожая картина наблюдается и в отношениях поставщик – продавец [10]. Все они работают с разными системами, а значит передача товарного контента между ними крайне затруднена. Разное программное обеспечение у участников рынка электронной коммерции вызвано разными задачами. Интернет-магазину важно создать витрину товаров, им более ценно получить УТП товаров (уникальное товарное предложение). Поставщикам и производителям важен технический учет товара и соответствие тактико-техническим характеристикам [11]. И часто, производственные системы производителей не умеют качественно сообщаться с другими системами внутри их компании, а значит и с покупателями, особенно если речь идет о международном сотрудничестве.

Однако есть попытки разработать нишевые стандарты для учета характеристик товаров. Широко известен стандарт ЕТИМ для электротехнических изделий. В России этот стандарт представлен в РАЭК¹. Подобные стандарты постепенно вводятся и в других отраслях, например медицине, химическом производстве, металлургии.

В статье представлен вариант формата, который позволяет участникам рынка электронной коммерции прийти к единому стандарту обмена данными. Данный формат можно рассматривать и как промежуточное средство обмена характеристиками товаров. Т. е. участникам обмена требуется провести всего одну синхронизацию – обработку описываемого формата. Однако отмечено, что у данного стандарта должен быть оператор для управления справочниками. Эта ниша бизнеса еще не занята.

¹ Единая база электротехнических товаров в международном стандарте ЕТИМ (Единая база РАЭК).

Цель – предложить формат описания и хранения характеристик товаров с сохранением логических связей между характеристиками.

Имеющиеся форматы обмена товарным контентом

Рынок электронной коммерции существует не один десяток лет. А значит участники рынка научились обмениваться данными о товарах. Ниже приведен не полный, но охватывающий самые распространенные форматы и стандарты для обмена информацией о товарах список:

1. **GDSN** (Global Data Synchronization Network) – это международный стандарт для синхронизации данных о товарах между производителями, дистрибьюторами и ритейлерами [13].

2. **Global Data Dictionary (GDD)** – это единый словарь данных, который используется для определения терминов, используемых в информационных системах, а также для их классификации и описания. GDD обеспечивает стандартизацию терминологии, что упрощает обмен данными между различными информационными системами, а также улучшает качество данных внутри организации.

3. **BMEcat** (Bundesverband Materialwirtschaft, Einkauf und Logistik e.V. catalog) – это стандартизированный формат обмена данными каталога на основе XML в сфере B2B

4. **Schema.org** – это коллективный проект от организаций Google, Microsoft, Yahoo и Yandex, который создан для улучшения способа представления структурированных данных в Интернете. Схема описывает набор типов данных и свойств, которые можно использовать для создания структурированных метаданных на веб-страницах. Она позволяет разработчикам сайтов и владельцам контента стандартизировать формат данных и сделать их более понятными для поисковых систем и других сервисов. Схема охватывает множество категорий, таких как события, организации, товары, места и другие, и может быть использована для описания любого типа контента в Интернете.

5. **ЕСЭТ** (Единая Система Электронной Торговли) – это система электронных закупок, созданная в России для автоматизации процесса закупок государственных и муниципальных организаций. Разработана в РАЭК (Ассоциация электронных коммуникаций).

6. **ETIM** (Electro Technical Information Model) – это стандарт, который разработан для описания электротехнических товаров. Он представляет собой структурированную базу данных, содержащую информацию о характеристиках товаров, таких как размеры, вес, цвет, энергопотребление и т. д. ETIM позволяет унифицировать описание товаров и обеспечивает эффективный обмен данными между производителями, поставщиками и покупателями. Стандарт ETIM используется в различных отраслях, включая электротехнику, строительство, промышленность и торговлю.

7. **GTIN** (Global Trade Item Number) – это международный стандарт идентификации товаров, который используется для присвоения уникальных идентификаторов товарам [13].

8. **EAN** (European Article Number) – это формат кодирования GTIN для европейских товаров.

9. **UPC** (Universal Product Code) – это формат кодирования GTIN для товаров, которые продаются в США и Канаде.

10. **ISBN** (International Standard Book Number) – это стандарт идентификации книг, который присваивается каждой новой книге и используется для уникальной идентификации и отслеживания ее продаж.

11. **ASIN** (Amazon Standard Identification Number) – это уникальный идентификатор, который присваивается каждому товару на платформе Amazon.

12. **XML** (eXtensible Markup Language) – это формат, который используется для хранения и обмена данными между различными системами и приложениями [14].

13. **YAML** (YAML Ain't Markup Language) – это формат сериализации данных, который используется для представления данных в человекочитаемой форме. YAML является языком разметки, который используется для создания конфигурационных файлов, сценариев и структурированных данных. YAML обычно используется для передачи данных между различными приложениями и языками программирования. Он отличается от других форматов сериализации, таких как JSON и XML, тем, что не использует закрывающие теги и вместо этого использует отступы для определения вложенности [15].

14. **CSV** (Comma-Separated Values) – это формат, который используется для хранения и передачи данных в виде таблицы, где значения разделены запятыми.

15. **JSON** (JavaScript Object Notation) – это формат, который используется для хранения и передачи данных в виде объектов JavaScript.

16. **YML** (Yandex Markup Language) – это язык разметки, который используется для описания структуры товарных каталогов, а также для передачи информации о товарах из интернет-магазина в Яндекс.Маркет и другие площадки.

17. Код стандартных продуктов и услуг Организации Объединенных Наций (**UNSPSC**) – это таксономия продуктов и услуги для использования в электронной коммерции. Это четырехуровневая иерархия, кодируемая как восьмизначное число, с необязательным пятым уровнем, добавляющим еще две цифры.

18. **TUN** – это датская стандартная система нумерации продуктов, идентифицирующая строительные материалы, управляемая (Trælasthandlerunionen). В настоящее время идентифицировано более 30 000 товаров. Номера TUN присваиваются поставщикам, идентифицирующим продукты, и поэтому продукт с несколькими поставщиками может иметь более одного номера TUN. Номера TUN в настоящее время сопоставляются с UNSPSC.

19. **eOTD** – это сокращение от открытого технического словаря ЕССМА. Словарь представляет собой независимую от языка базу данных концепций со связанными терминами, определениями и изображениями, используемыми для однозначного описания людей, организаций, местоположений, товаров, услуг, процессов, правил и положений. EOTD поддерживается Ассоциацией управления кодами электронной торговли (ЕССМА).

20. **Icescat** – это открытый каталог продуктов, содержащий информацию о более чем 1 миллионе продуктов от более чем 14000 производителей. Он предоставляет бесплатный доступ к информации о продуктах, включая технические характеристики, изображения, видео и маркетинговые материалы, для использования в электронной коммерции, маркетинге и продажах. Icescat основан на открытых стандартах, таких как ЕТИМ, UNSPSC и Open Icescat, и обеспечивает единый формат данных для удобства использования в различных системах.

21. **ISO 22745** – это стандарт, который устанавливает методы и принципы описания характеристик товаров и услуг с использованием языка маркировки данных XML. Он включает в себя набор правил и рекомендаций по описанию свойств продуктов, включая физические, химические, технические и коммерческие характеристики. Стандарт ISO 22745 разработан для облегчения обмена информацией о товарах и услугах между компаниями и организациями.

22. **ePASS** является централизованным белорусским информационным ресурсом, содержащим описания товаров в формате, соответствующем международным стандартам электронной торговли.

Наличие такого внушительного числа форматов явно говорит о следующем:

1. Рынок все еще не пришел к единому стандарту.
2. Рынок не останавливает попыток прийти к единому стандарту.

Однако практика работы интернет-магазинов в России с их поставщиками показывает, что зачастую обмена контентом либо нет вообще, либо он осуществляется в формате текстовых или графических файлов, т. е. форматами без возможности автоматического чтения и разбора данных (автоматический сбор и распознавание контента не рассматривается ввиду высокой стоимости этого процесса и очень высокого риска получить ошибки).

Самым распространенным в России является формат YML. Подробная документация по нему опубликована на сайте Яндекса. Из публичных форматов обмена характеристиками, это единственный формат, позволяющий проводить такой обмен. Международные форматы пока не получили своего широкого применения в России.

Вот пример передачи характеристик товара по формату YML:

```
<param name="Объем (л)" code="ob_em_l">1.7</param>  
<param name="Гарантия" code="garantiya">1 год</param>
```

Как видно, и характеристика, и ее значение передаются простым текстовым значением. Такой формат данных крайне сложно импортировать в виду необходимости настраивать большой объем справочников. Формат очень наглядный для визуального осмотра, но не подходит для автоматического обмена.

Разработка полной документации на все типы товаров, все рынки и направления – крайне объемная работа, которую не выполнить без привлечения участников рынка к общему диалогу. Самым критическим элементом предлагаемого формата является общий реестр полей (их названия и идентификаторы), для выработки которого все участники рынка (или сегмента рынка) должны совместно его обсудить. А в идеальном варианте создать единого оператора данных, который будет агрегировать на себе все данные, поддерживать их актуальность и оперативно обновлять.

Связи характеристик

В подавляющем большинстве систем управления контентом о товарах реализована иерархическая структура зависимостей данных. На верхнем уровне идут шаблоны товаров, далее – группы характеристик, далее – сами характеристики. Часто, для характеристики внутри шаблона используется свой справочник значений.

В более продвинутых системах предусмотрен сквозной справочник значений. Например, справочник «цвет» разумно применять одновременно на все типы товаров.

Но дальше, чем применение сквозных справочников, мало кто ушел. Вызвано это больше отсутствием в широком доступе соответствующего уровня ПО (готовых решений). А переделывать уже существующее решения часто не целесообразно по бизнес показателям. Однако запрос на подобное усовершенствование постепенно формируется. К примеру, за последние 3-5 лет в России активно начали развиваться PIM-системы, что явно говорит о росте спроса на качественные системы управления контентом.

Словари. Иначе их еще называют «справочники». Их можно разделить на три типа:

1. Классический. Просто набор значений в виде списка.
2. Вертикально зависимый. В нем допускается связь на уровне «совместимость» между разными строками из этого словаря. Например, берем словарь «Версия USB». Зная саму технологию USB, можно утверждать, что для товаров, поддерживающих USB 3.0, одновременно работает и поддержка USB 2.0.

3. Горизонтально зависимый. Наличие возможных значений текущего справочника зависят от значений другого справочника. Например, справочник «Версия USB» и «Скорость записи данных». Если в товаре указали некую версию USB, то очевидно, можно соответствующим образом сузить варианты из справочника скорость записи данных.

Допускается наличие многоуровневых горизонтальных зависимостей, а также комбинаций с вертикальными зависимостями.

При записи в базу данных значения для товара, стоит сразу же записывать и все вертикальные и горизонтальные связи, даже если их не указал пользователь системы. Это позволит реализовать все необходимые фильтры на уровне запросов в базу данных без дополнительных пересчетов. Однако есть нюанс, связанный с изменением текущих зависимостей и появлением новых с течением времени. Будет верным, при появлении нового значения тут же обновить данные по всем товарам. Но если в базе миллионы товаров, то такое обновление потребует времени. Плюс высокий риск получить ошибку. Поэтому архитекторам такой базы данных рекомендуется самим определить, смогут ли они вносить изменения, или же лучше применить подход динамического расчета зависимостей при запросе в базу данных с учетом правил зависимости внутри словарей.

Для числовых значений и в словарях, и в отдельных полях (характеристиках) рекомендуется все данные хранить в приведенном виде. Речь идет о приведении всех единиц измерения к единому значению. Это позволяет динамически сравнивать значения при обращении к базе данных. При этом для вывода этих значений в запросе можно (нужно) применять конвертацию в зависимости от требования шаблона товара. Например, рассмотрим параметр «вес товара» (точнее будет сказать «масса», однако в общей практике потребители привыкли к слову «вес»). Для товара «гречка» логично, чтобы масса измерялась в килограммах. А вот для товара «мобильный телефон» уже применяется масса в граммах. Как видно, тип данных одинаковый, хранить в базе данных это стоит также единообразно, а вот перед выводом данных на клиента данные стоит конвертировать для понятного вида. Какую именно единицу измерения брать за основу для хранения в базе данных – выбор для авторов системы.

Шаблоны товаров. От классического понятия «шаблон» рекомендуется отказаться. Его роль должна заключаться лишь в представлении вида товара для потребителя, т. е. только в визуализации.

Это реализуется за счет того, что понятие «характеристика» должно стать самостоятельной сущностью, при этом единой и сквозной на все товары. Иными словами, не допускается, чтобы в целом в базе данных было две одинаковые характеристики как сущности. Если есть два типа товаров, которые используют одну и ту же характеристику (например, цвет), то шаблоны этих типов товаров должны ссылаться на одну эту характеристику.

Это позволит реализовать систему поиска, фильтрации и сравнения товаров вне зависимости от их типа. Например, можно будет сравнить качество экранов для телефонов, телевизоров, видеодомофонов, МФУ (где есть экраны) и т. д. Как видно, гибкость в возможностях сравнения становится безграничной, по сравнению с классической иерархической структурой шаблонов.

Шаблоны стоит делать с возможностью ссылаться на другие шаблоны. Например, для товара «Сервер» нужен свой соответствующий шаблон. Но по факту, сервер состоит из разных компонентов, каждый из которых является самостоятельным товаром и имеет свой шаблон. Т.о. создавая шаблон для сервера, логично просто ссылаться на другие шаблоны, объединить их.

Базовая структура формата

Чтобы передать контент точно и полно, все данные следует разделить на отдельные сущности: шаблон, словарь, характеристика, товар, константа. Это позволяет значительно уменьшить размер (в байтах) передаваемого контента по конкретному товару, плюс позволяет автоматизировать процессы расширения базы данных получателя контента.

В качестве базового формата предлагается выбрать XML с кодировкой UTF-8, а на его основе уже сделать вторичные варианты на базе форматов JSON или JSONP. Допускается работа и с другими кодировками, что особенно актуально на международном уровне.

XML позволяет говорить о низком пороге входа формата, он известен и широко применяется. Существуют множество готовых скриптов и алгоритмов для автоматического разбора данных из XML, в том числе на уровне баз данных.

JSON позволяет реализовать более легкий обмен с web-приложениями без потери полноты и качества передаваемых данных. Данный формат все активнее вытесняет XML, хотя он все же проигрывает ему в ряде особенностей. Например, JSON не позволяет передавать атрибуты.

В качестве стандарта для наименований полей рекомендуется опираться на schema.org, например, /Product для полей о товаре.

На первом уровне вложения рекомендуется создавать общие группы для контента. Это позволит далее наращивать предлагаемый формат новыми типами данных, а также обеспечивать особенности данных по сегментам рынка (типам товаров). Вместе с названием группы вводить и атрибут версии формата для этой группы. Это позволит развивать формат внутри конкретной группы, не затрагивая изменения в других группах данных. Пример:

```
<?xml version="1.0" encoding="UTF-8"?>
<pml date="2023-01-01T12:40:40+01:00" lang="ru">
  <goods version="1.0">
    <good id="242432" template="33">
      <feature id="55334">1.7</feature>
      <feature id="5474">
        <value>23</value>
        <value>55</value>
      </feature>
    </good>
  </goods>
</pml>
```

- *pml* – это Product Markup Language (по аналогии с XML);
- *lang* – локализация ответов, по умолчанию следует использовать EN;
- *goods* – это группа для описания товаров;
- *version* – это версия формата вложенных данных.

Параметр *feature* содержит в себе только ссылку на характеристику в шаблоне. Чтение данных этого поля стоит рассматривать в контексте шаблона. Если в шаблоне эта характеристика имеет простое значение (не словарь), то данные этой характеристики интерпретируются буквально, как числовое значение (или текстовое). Если же шаблон говорит, что он тут ожидает увидеть *id* элемента словаря, или несколько таких *id*, то данные характеристики уже стоит интерпретировать как *id*. Если ссылок на словарь несколько, то они передаются как *value* внутри *feature*. Аналогичный подход применяется и к прямым данным, когда их несколько. Например, если речь про

дискретный тип характеристики без ссылок на словари (комплект поставки, где каждый элемент записывается текстовой строкой).

По аналогии решается ситуация с записью таких типов характеристик, как диапазон значений, где вместо *value* нужно указать *min* и *max* в виде значения или ссылки на словарь.

Схема, при которой в *feature* передаются только идентификаторы и значения, позволяет значительно сократить размер файла и минимизировать вероятность ошибок при неверном кодировании данных. Но для целей визуальной проверки полноты и точности данных допускается обогатить данные соответствующими атрибутами для передачи непосредственно значений. Тогда примерный вид будет такой:

```
<?xml version="1.0" encoding="UTF-8"?>
<pml date="2023-01-01T12:40:40+01:00" lang="ru">
  <goods version="1.0">
    <good id="242432" template="33">
      <feature
        feature_id="55334"
        name="Screen diagonal (inch)"
        name_ru="Диагональ экрана (дюйм)">
        1.7
      </feature>
      <feature feature_id="5474" name="Color" name_ru="Цвет">
        <value name="Red" name_ru="Красный">23</value>
        <value name="White" name_ru="Белый">55</value>
      </feature>
    </good>
  </goods>
</pml>
```

Если характеристика передана с единицей измерения, то ее указываем в скобках вместе с названием характеристики. Выделять под единицу измерения дополнительный параметр не имеет смысла, т. к. для целей машинной обработки есть поля описания шаблонов. А в указанном примере вывода цель – визуальное ознакомление.

Для передачи данных о словарях создается группа *dictionaries*.

```
<?xml version="1.0" encoding="UTF-8"?>
<pml date="2023-01-01T12:40:40+01:00" lang="ru">
  <dictionaries version="1.0">
    <dictionary id="5474" name="Color" name_ru="Цвет">
      <property id="23" value="Red">Красный</property>
      <property id="55" value="White">Белый</property>
    </dictionary>
  </dictionaries>
</pml>
```

Как можно заметить, здесь не указываются элементы зависимости значений словарей. Подразумевается, что все зависимости уже выставлены в виде прямых связей товара на значения словарей. Иными словами, зависимости должны быть настроены в той системе, которая имеет интерфейсы создания нового контента, а для целей передачи этого контента потребителю такие зависимости уже не требуются.

Значение для полей *property* передается с учетом запрашиваемой локализации. При этом в атрибуте *value* всегда стоит передавать значение поля на EN локализации.

Допускается не передавать *value*, если запрашиваемая локализация соответствует умолчанию. Данный подход позволяет стандартизировать передачу значений.

Для описания характеристик как составных частей шаблона вводиться группа *features*:

```
<?xml version="1.0" encoding="UTF-8"?>
<pml date="2023-01-01T12:40:40+01:00" lang="ru">
  <features version="1.0">
    <feature dictionary_id="5474" name="Color" name_ru="Цвет">
      <type id="5474">12</type>
    </feature>
  </features>
</pml>
```

Название характеристики выводится сразу, без дополнительного запроса. Типы характеристик нужно реализовать как расширяемый список, что позволит учесть особенности товаров и их свойств из самых разных отраслей. Если характеристика ссылается на словарь, то вводиться атрибут *dictionary_id*. Указание типа характеристики делается путем указания идентификатора из словаря типов характеристик. Сам словарь является некой константой, который передавать внутри файла не имеет смысла, его следует сделать публичным в документации к формату. Ниже приведен список самых распространенных типов характеристик:

- *dictionary* – одно значение из словаря;
- *multy_dictionary* – множественное значение из словаря;
- *bool* – булево значение true / false;
- *number* – числовое значение, включая дробные;
- *text* – текстовое значение;
- *link / multy_link* – ссылка / ссылки;
- *img / multy_link* – хранение картинки внутри базы данных (редко используется, как правило картинки хранят в виде ссылок);
- *file / files* – ссылки на электронные документы;
- *interval* – диапазон числовых значений от минимального до максимального (не дискретный, для дискретного интервала как правило применяют *multy_dictionary*);
- *html* – текстовое поле для хранения текста с учетом *html* разметки;
- *size* – спец характеристика для записи габаритов в виде самостоятельных полей высоты, ширины и глубины.

Для расширения описательных возможностей поля *type* можно (нужно) вводить дополнительные атрибуты к этому полю, которые будут зависеть от указанного типа поля *type*. Вот пример таких атрибутов, но стоит понимать, что их список является расширяемым:

- *max / min* – обозначение предельных значений для числовых форматов, длины текста для текстовых, а также для множественного выбора из словарей;
- *name* – для ввода имени, например, для типов *link, file, img* и т. п.;
- *mask* – указание формата строки или числа, валидация поля;
- *unit* – для указания единицы измерения (для этого следует объявить отдельный словарь в документации), используется как базовая единица измерения, в которой данные хранятся в базе данных.

Для сущности «шаблон» вводиться группа *templates*, внутри которого создаются группы характеристик.

```
<?xml version="1.0" encoding="UTF-8"?>
<pml date="2023-01-01T12:40:40+01:00" lang="ru">
```

```

<templates version="1.0">
  <template id="242432" name="Phone" name_ru="Телефон">
    <group id="24323" name="Display" name_ru="Экран" sort="4"/>
      <feature
        id="55334"
        feature_id="55334"
        rank="3"
        unit="66"
        sort="4"/>
    </group>
  <group template_id="333"></group>
</template>
</templates>
</pml>

```

Название шаблона и групп выводится всегда с учетом локализации. Идентификатор в *feature* отвечает за непосредственную идентификацию характеристики внутри группы и шаблона. За привязку характеристики к ее описанию отвечает *feature_id*. Далее добавляются атрибуты, которые «настраивают» характеристику для конкретно этого шаблона (опять же, атрибуты расширяемые):

1) *sort* (integer) – сортировка характеристики или группы, чем меньше число, тем важнее характеристика;

2) *unit* – единица измерения, она может отличаться от единицы измерения в описании характеристики, итоговое значение этой характеристики в данных о товаре уже должны выводиться с учетом настройки шаблона;

3) *rank* – тип визуализации характеристики, как правило, есть 4 типа:

– базовая, т. е. без которой описать товар невозможно, например, диаметр сечения для кабеля;

– обязательная, требующая обязательного заполнения редактором и которая клиента интересуют в первую очередь при ознакомлении с товаром;

– вторичная, все прочие характеристики, нужные для общего описания товара;

– скрытые, характеристики, не важные для клиента, но требующиеся для некой автоматизации.

Для организации вложенного шаблона нужно для *group* ввести атрибут *template_id*, что и укажет на вложенность. Далее, можно вводить атрибуты и вложенные элементы для переопределения некоторых значений, если это требуется. Например, возьмем шаблон «дисплей», у которого есть характеристика «число цветов». Для мобильных телефонов эта характеристика давно перестала быть актуальной для пользователя, т. е. ей лучше поставить *rank="вторичная"*. А вот для дисплея МФУ эта характеристика уже может иметь интерес и там можно переопределить *rank="обязательная"*.

Заключение

Предлагаемый формат описания товара и его характеристик основан на опыте работы нескольких интернет-магазинов Группы компаний Икс-Ком и коммуникации с их партнерами, как поставщиками, так и потребителями.

Как видно из описания формата, он предельно гибкий на расширение, позволяет обеспечить передачу всех типов данных, связать их между собой. Все данные в нем передаются в виде словарей путем указания их идентификаторов, что позволяет значительно снизить риск искажения данных при их пересылке и декодировании, а также

упростить автоматическое чтение. Также такой подход позволяет реализовать чтение / запись напрямую в базу данных, соблюдая реляционный подход распределения данных по таблицам.

Применение подхода централизованного управления словарями, названиями, константами позволит рынку или сегментам рынка прийти к единому стандарту взаимоотношений.

Применение подхода расширения словарей путем установки зависимостей позволяет реализовать обратно совместимое расширение данных. Например, у товара «USB удлинитель» в качестве указания на тип гнездо / штекер будет простая ссылка на словарь форматов, т. е. на «USB». И в целом этого часто достаточно, чтобы указать это на сайтах. А вот указание гнездо или штекер можно организовать как зависимый справочник. В итоге, те потребители, кому важно получить подробные данные, смогут это сделать. При этом все другие потребители продолжают получать характеристику в первичном виде, не обращая внимание на расширение через зависимости.

СПИСОК ИСТОЧНИКОВ

1. Obasun O. Exploring the Utilization of Internet Technology for Creating a New Industrial Order. *International Journal of Innovative Science and Research Technology*. 2023;6(1):1682–1703. DOI: 10.5281/zenodo.7620863.
2. Челябинка В., Лизакова Р.А. Особенности продажи товаров на маркетплейсах. *Умная цифровая экономика*. 2022;2(3):12–16.
3. Соболевский Ю.С. Визуальный и текстовый контент как способ продвижения товаров на «Яндекс.Маркете» на примере рынка электроинструментов. *Экономика XXI века – экономика новых возможностей. Актуальные вопросы теории и практики экономики и управления: сборник научных статей*. СПб.: Ассоциация НИЦ Пересвет; 2021. С. 52–57.
4. Akram U., Hui P., Kaleem Khan M., Tanveer Y., Mehmood K., How website quality affects online impulse buying: Moderating effects of sales promotion and credit card use. *Asia Pacific Journal of Marketing and Logistics*. 2018;30(1):235–256. DOI: 10.1108/APJML-04-2017-0073.
5. Кокшаров С.П., Козубня А.Е. Создание страницы товара в интернет–магазине. *Инновационное развитие экономики: российский и зарубежный опыт: сборник статей по итогам Международной научно-практической конференции*. Стерлитамак: Агентство международных исследований; 2021. С. 67.
6. Шлеина В.В. Проблемы оценки характеристик товаров. *Проблемы идентификации, качества и конкурентоспособности потребительских товаров: сборник трудов конференции*. Курск: ЗАО «Университетская книга»; 2018. С. 440–443.
7. Bednár P., Delina R. Towards the Semantic Platform for Digital Single Market. In: *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*. Košice, Slovakia; 2018. p. 165–168. DOI: 10.1109/DISA.2018.8490616.
8. Горчатов О.В., Губин В.В., Нечипоренко П.Р. Единая система электронной паспортизации продукции на основе стандартизованного реестра параметров. *ИТ-Стандарт*. 2018;15(2):18–21.
9. Battistello L., Haug A., Suzic N., Hvam L. Implementation of product information management systems: Identifying the challenges of the scoping phase. *Computers in Industry*. 2021;133:103533. DOI: 10.1016/j.compind.2021.103533.
10. Яковлев А.А., Глухов В.В. Модель взаимодействия производителей и потребителей через основные характеристики товара: полезность, качество, цена, меновая

- стоимость. *Научно-технические ведомости СПбГПУ. Экономические науки*. 2018;11(3):194–202. DOI: 10.18721/JE.11317.
11. Долганов К.Б., Бром А.Е. Разработка методики проектирования складских комплексов в условиях существенной неопределенности массово-габаритных характеристик товаров. *Подъемно-транспортное дело*. 2015;82(4-5):41–43.
 12. Chen H. C., Prater E. Information system costs of utilizing electronic product codes in achieving global data synchronization within the pharmaceutical supply chain network. *International Journal of Information Systems and Supply Chain Management (IJISSCM)*. 2013;6(1):62–76. DOI: 10.4018/jisscm.2013010104.
 13. Brock D.L. Integrating the electronic product code (EPC) and the global trade item number (GTIN). *White Paper*. 2001;25:2–25. URL: https://cocoa.ethz.ch/downloads/2014/06/None_MIT-AUTOID-WH-004.pdf
 14. Bray T., Paoli J., Sperberg-McQueen C.M., Maler E., Yergeau F. Extensible markup language (XML). *World Wide Web Journal*. 1997;2(4):29–66.
 15. Ben-Kiki O., Evans C., Ingerson B. Yaml ain't markup language (yaml™) version 1.1. *Working Draft*. 2009;5:11.

REFERENCES

1. Obasun O. Exploring the Utilization of Internet Technology for Creating a New Industrial Order. *International Journal of Innovative Science and Research Technology*. 2023;6(1):1682–1703. DOI: 10.5281/zenodo.7620863.
2. Chelyapina V., Lizakova R.A. Features of selling goods on marketplaces. *Umnaya tsifrovaya ekonomika = Smart digital economy*. 2022;2(3):12–16. (In Russ.).
3. Sobolevsky Y.S. Visual and textual content as a way to promote goods on «Yandex.Market» on the example of the market of power tools. In: *Economics of the XXI century – the economy of new opportunities. Topical issues of theory and practice of economics and management*. Saint-Petersburg, Association RC Peresvet; 2021. p. 52–57. (In Russ.).
4. Akram U., Hui P., Kaleem Khan M., Tanveer Y., Mehmood K., How website quality affects online impulse buying: Moderating effects of sales promotion and credit card use. *Asia Pacific Journal of Marketing and Logistics*. 2018;30(1):235–256. DOI: 10.1108/APJML-04-2017-0073.
5. Koksharov S.P., Kozubnya A.E. Creating a product page in an online store. In: *Innovative development of the economy: Russian and foreign experience: Proceedings of International Scientific Conference*. Sterlitamak, AMI; 2021. p. 67. (In Russ.).
6. Shleina V.V. Problems of assessing the characteristics of goods. In: *Problems of identification, quality and competitiveness of consumer goods*. Kursk, JSC Universitetskaya kniga; 2018. p. 440–443. (In Russ.).
7. Bednár P., Delina R. Towards the Semantic Platform for Digital Single Market. *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*. Košice, Slovakia; 2018. p. 165–168. DOI: 10.1109/DISA.2018.8490616.
8. Gorchatov O.V., Gubin V.V., Nechiporenko P.R. A unified system for electronic certification of products based on a standardized register of parameters. *IT-Standard = IT-Standard*. 2018;15(2):18–21. (In Russ.).
9. Battistello L., Haug A., Suzic N., Hvam L. Implementation of product information management systems: Identifying the challenges of the scoping phase. *Computers in Industry*. 2021;133:103533. DOI: 10.1016/j.compind.2021.103533.
10. Yakovlev A.A., Glukhov V.V. Model of interaction between producers and consumers through the main characteristics of the goods: usefulness, quality, price, exchange value.

- Nauchno-tekhnicheskie vedomosti SPbGPU. Ekonomicheskie nauki = Scientific and technical statements of SPbSPU. Economic Sciences.* 2018;11(3):194–202. DOI: 10.18721/JE.11317.
11. Dolganov K.B., Brom A.E. Development of a methodology for designing warehouse complexes in conditions of significant uncertainty in the mass-dimensional characteristics of goods. *Pod'emno-transportnoe delo = Lifting and transport business.* 2015;82(4-5):41–43. (In Russ.).
 12. Chen H. C., Prater E. Information system costs of utilizing electronic product codes in achieving global data synchronization within the pharmaceutical supply chain network. *International Journal of Information Systems and Supply Chain Management (IJISSCM).* 2013;6(1):62–76. DOI: 10.4018/jisscm.2013010104.
 13. Brock D.L. Integrating the electronic product code (EPC) and the global trade item number (GTIN). *White Paper.* 2001:25:2–25. URL: https://cocoa.ethz.ch/downloads/2014/06/None_MIT-AUTOID-WH-004.pdf
 14. Bray T., Paoli J., Sperberg-McQueen C.M., Maler E., Yergeau F. Extensible markup language (XML). *World Wide Web Journal.* 1997;2(4):29–66.
 15. Ben-Kiki O., Evans C., Ingerson B. Yaml ain't markup language (yaml™) version 1.1. *Working Draft.* 2009;5:11.

ИНФОРМАЦИЯ ОБ АВТОРЕ / INFORMATION ABOUT THE AUTHOR

Шишкин Андрей Геннадьевич, Andrey Gennadievich Shishkin, Chief technical director, группа компаний Икс-Ком, Москва, Российская Федерация. Technical Officer at X-Com Group of Companies, Moscow, the Russian Federation.
e-mail: andrey.shishkin.spb@gmail.com
ORCID: [0009-0006-6983-1939](https://orcid.org/0009-0006-6983-1939)

Статья поступила в редакцию 01.06.2023; одобрена после рецензирования 06.07.2023; принята к публикации 26.07.2023.

The article was submitted 01.06.2023; approved after reviewing 06.07.2023; accepted for publication 26.07.2023.