R

WIPO
ВСЕМИРНАЯ ОРГАНИЗАЦИЯ
ИНТЕЛЛЕКТУАЛЬНОЙ
СОБСТВЕННОСТИ

CWS/14/24
ОРИГИНАЛ: АНГЛИЙСКИЙ
ДАТА: 6 ОКТЯБРЯ 2026 ГОДА

Комитет по стандартам ВОИС (КСВ)

Четырнадцатая сессия
Женева, 23–27 ноября 2026 года

ПРЕДЛАГАЕМЫЙ НОВЫЙ СТАНДАРТ ВОИС ДЛЯ ПОДДЕРЖКИ РАБОТЫ В ОБЛАСТИ ОЧИСТКИ ДАННЫХ ОБ ИМЕНАХ

Документ подготовлен руководителями Целевой группы по стандартизации имен

РЕЗЮМЕ

1. Целевая группа по стандартизации имен представляет окончательный проект нового стандарта ВОИС по очистке данных об именах для рассмотрения и утверждения на четырнадцатой сессии Комитета по стандартам ВОИС (КСВ или Комитет).

СПРАВОЧНАЯ ИНФОРМАЦИЯ

2. В 2023 году на одиннадцатой сессии КСВ утвердил измененную формулировку задачи № 55 в следующей редакции:

«Подготовить предложение, касающееся дальнейших действий по достижению стандартизации имен в документах по интеллектуальной собственности (ИС), в целях разработки стандарта ВОИС для помощи ведомствам ИС в обеспечении более высокого качества исходной информации, касающейся имен»

(см. пункты 75–78 документа CWS/11/28).

3. На той же сессии в 2023 году КСВ рассмотрел новый перечень руководящих указаний, предложенных Целевой группой по стандартизации имен для поддержки работы в области очистки данных об именах заявителей. Комитет договорился использовать формулировку «рекомендации» вместо «руководящие указания» в названии нового предлагаемого стандарта. КСВ также отметил, что Секретариат предложил назвать рекомендации по очистке данных об именах «стандартом ВОИС ST.93» (см. пункт 135 документа CWS/11/28).

4. В то же время КСВ не утвердил предложенный стандарт и вернул его в Целевую группу по стандартизации имен для дальнейшего обсуждения и доработки. Более того, Комитет принял к сведению, что Секретариат планирует изучить возможность размещения набора таблиц транслитерации на сайте ВОИС (см. пункты 136 и 137 документа CWS/11/28).

5. В 2024 году на двенадцатой сессии КСВ рассмотрел доработанный проект предлагаемого стандарта ВОИС по очистке данных об именах, представленный Целевой группой по стандартизации имен. Предложенный стандарт поддержали несколько делегаций. Однако КСВ не утвердил его, поскольку одна из делегаций запросила дополнительное время для тщательной оценки потенциального влияния внедрения предлагаемого стандарта и для проведения внутренних консультаций и консультаций с клиентами. Комитет поручил Целевой группе пересмотреть проект стандарта и продолжить усилия по его совершенствованию по мере необходимости (см. пункты 85–91 документа CWS/12/29).

6. На той же сессии в 2024 году КСВ поручил Международному бюро организовать практикум по теме очистки данных об именах в 2025 году с участием всех желающих. КСВ также призвал своих членов и наблюдателей оказать Международному бюро поддержку в вопросе проведения этого мероприятия и активно распространять информацию о нем и участвовать самим (см. пункт 92 документа CWS/12/29).

7. На той же двенадцатой сессии КСВ принял к сведению, что стандарты ВОИС представляют собой рекомендации, основанные на передовом опыте. В общем плане ведомства ИС и представители отрасли внедряют стандарты ВОИС по мере готовности к этому процессу и с учетом своих потребностей. В исключительных случаях ведомства ИС договариваются внедрить тот или иной утвержденный стандарт в одни и те же сроки (как в случае со стандартом ВОИС ST.26) или в согласованном порядке (как в случае со стандартом ВОИС ST.92). В дальнейшем все стандарты ВОИС могут быть доработаны с учетом опыта внедрения в ведомствах ИС или новых потребностей: это обеспечивает их неизменную практическую ценность, эффективность и способность адекватно реагировать на меняющиеся требования. КСВ также принял к сведению, что предлагаемый стандарт ВОИС ST.93 носит такой же общий характер, как и другие стандарты ВОИС, и ведомства ИС могут внедрить его сразу после принятия, внедрять поэтапно или не внедрять в принципе, если сочтут, что действующие системы адекватны. ST.93, как и любой другой стандарт ВОИС, может быть впоследствии усовершенствован с учетом опыта внедрения в ведомствах ИС.

8. В соответствии с решением, принятым Комитетом на двенадцатой сессии, 12 мая 2025 года Международное бюро организовало практикум по стандартизации имен. На заседании Целевой группы по стандартизации имен, состоявшемся 13 мая 2025 года, были проанализированы итоги этого мероприятия и проведено заключительное обсуждение в преддверии подготовки окончательного проекта рекомендаций по очистке данных об именах. Подробная информация о работе Целевой группы и прогрессе, достигнутом в межсессионный период, представлена в документе CWS/13/7.

9. Что касается таблиц транслитерации, которыми пользуются ведомства ИС, Целевая группа проинформировала участников двенадцатой сессии КСВ о том, что всем членам Группы было предложено направить в Секретариат имеющиеся у них таблицы. Это позволит пользователям и ведомствам ИС знакомиться с таблицами других ведомств, работающих на разных языках и с разными системами письма, тем самым обеспечивая единообразие и эффективность коммуникации между ведомствами и их клиентами и не вынуждая ведомства вносить изменения в действующие базы данных. Для того чтобы это стало возможным, ведомствам ИС предлагается поделиться своими таблицами, если такие существуют.

10. На тринадцатой сессии КСВ рассмотрел доработанный проект предлагаемого стандарта ВОИС по очистке данных об именах (ST.93), но не утвердил его в силу отсутствия консенсуса. КСВ поручил Целевой группе по стандартизации имен пересмотреть и дополнительно доработать проект стандарта ВОИС ST.93 и представить окончательное предложение для рассмотрения на следующей сессии Комитета (см. пункты 105 и 106 документа CWS/13/33).

ПРЕДЛАГАЕМЫЙ НОВЫЙ СТАНДАРТ

Цели

11. Цель предлагаемого стандарта — предложить ведомствам ИС рекомендации общего характера по очистке данных об именах заявителей с целью повышения единообразия, точности и достоверности таких данных. Предлагаемый стандарт призван помочь в поиске членов семейств патентов-аналогов, проведении статистического анализа и использовании информации об ИС для других целей.

Преимущества

12. Предлагаемый стандарт позволит ведомствам ИС подходить к вопросу повышения качества, единообразия и удобства использования данных об именах заявителей с общих позиций. Его внедрение упростит поиск и сопоставление данных, относящихся к разным массивам, юрисдикциям и временным интервалам, и тем самым обеспечит более качественный поиск информации об ИС, управление портфелем прав, статистический анализ и обмен данными. Он также поможет уменьшить дублирование данных и устранить нестыковки в них, будет способствовать автоматизации процедур обработки данных, созданию инструментов, ориентированных на работу с данными, и большей унификации подходов при решении вопросов, связанных с разным написание имен, сокращенными названиями и другими различиями, вызванными использованием разных языков и систем письма.

Сфера охвата

13. В предлагаемом стандарте приводятся общие рекомендации для ведомств ИС по получению, обработке, очистке и публикации данных об именах заявителей. Стандарт не охватывает конкретные способы и технические решения очистки данных, адаптации имен или их передачи на других языках, включая транслитерацию, транскрибирование и перевод. Более того, он не предлагает конкретных методов стандартизации имен, будь то выбор алгоритмов, указание этапа, на котором необходимо проводить преобразование, частота преобразований или стратегии объединения записей.

Изменения, внесенные в последний проект

14. В русле решения, принятого на тринадцатой сессии КСВ, членам Целевой группы по стандартизации имен было предложено повторно изучить и прокомментировать представленный на сессии проект стандарта ВОИС ST.93. С учетом озвученных замечаний в проект были внесены следующие изменения:

- в пункт 3 добавлены определения терминов «валидация» и «устранение неоднозначности»; и

- формулировка пункта 11 скорректирована таким образом, чтобы в ней проводилось четкое различие между языками и наборами символов.

15. Целевая группа также изучила предложение включить в данный стандарт в порядке исключения оговорку о добровольном характере предлагаемых рекомендаций. В свете дискуссии Целевой группы и единообразия с другими стандартами ВОИС Секретариат предлагает включить (в виде исключения) в стандарт редакционное примечание Международного бюро следующего содержания:

«Внедрение настоящего стандарта является добровольным и факультативным. Каждое ведомство самостоятельно, исходя из конкретных обстоятельств, определяет, следует ли ему применять настоящий стандарт, когда и в каком объеме — полностью, частично или отказаться от его применения».

16. Предлагаемый текст стандарта приводится в приложении к настоящему документу. Все изменения, внесенные в текст, представленный на тринадцатой сессии КСВ, показаны в режиме правки: предлагаемые исключения вычеркнуты, новые формулировки — подчеркнуты.

17. Предлагается следующее название нового стандарта ВОИС:

«Стандарт ВОИС ST.93. Рекомендации по очистке данных об именах».

18. Если Комитет утвердит новый стандарт на текущей сессии, предлагается поручить Секретариату опубликовать его в части 3 Справочника ВОИС по информации и документации в области интеллектуальной собственности.

ПРЕДЛОЖЕНИЕ ПО ПЕРЕСМОТРУ ЗАДАЧИ № 55

19. Если предлагаемый новый стандарт ВОИС ST.93 будет утвержден КСВ, работу по выполнению задачи № 55 в текущей редакции можно будет считать выполненной. Целевой группе предлагается продолжить свою деятельность с целью поддержки внедрения нового стандарта ведомствами ИС и сохранения его актуальности и последующего совершенствования с учетом опыта реализации. Дальнейший сбор информации о практике очистки данных об именах клиентов и обмен этими сведениями станут отличной основой для этой работы. В связи с этим руководители Целевой группы предлагают изменить формулировку задачи № 55 следующим образом:

«Обмениваться практическим опытом в области очистки данных об именах клиентов, включая используемые алгоритмы и цели, для которых применяются очищенные или унифицированные данные; разрабатывать практические рекомендации для ведомств ИС; поддерживать стандарт ВОИС ST.93 в актуальном состоянии и при необходимости обновлять его».

ПРЕДЛАГАЕМАЯ ПУБЛИКАЦИЯ В ЧАСТИ 7 СПРАВОЧНИКА ВОИС ТАБЛИЦ ТРАНСЛИТЕРАЦИИ, ИСПОЛЬЗУЕМЫХ ВЕДОМСТВАМИ ИС

20. С учетом обсуждения преимуществ использования ведомственных таблиц транслитерации имен клиентов предлагается собрать и опубликовать эти материалы на сайте ВОИС в части 7 Справочника ВОИС по информации и документации в области интеллектуальной собственности. Возможно, наличие такого централизованного ресурса будет способствовать точному сопоставлению и проверке имен клиентов в контексте разных юрисдикций.

21. Предлагается рекомендовать ведомствам ИС представить в Секретариат URL-ссылки на используемые ими таблицы транслитерации для публикации в части 7 Справочника ВОИС.

22. Секретариат предлагает добавить в часть 7 Справочника ВОИС новый подраздел «Используемые ведомствами ИС таблицы транслитерации», в котором будут публиковаться представленные материалы.

23. КСВ предлагается:

(a) принять к сведению информацию, изложенную в настоящем документе и приложении к нему;

(b) рассмотреть и утвердить включение в предлагаемый новый стандарт ВОИС редакционного примечания, упомянутого в пункте 15 выше;

(c) рассмотреть и утвердить название нового стандарта ВОИС, упомянутое в пункте 17 выше;

(d) рассмотреть и утвердить новый стандарт ВОИС ST.93, упомянутый в пунктах 11–16 выше и представленный в приложении к настоящему документу;

(e) поручить Секретариату опубликовать новый стандарт ВОИС ST.93 в части 3 Справочника ВОИС, как указано в пункте 18 выше;

(f) рассмотреть и утвердить новую предлагаемую формулировку задачи № 55, упомянутую в пункте 19 выше; и

(g) поручить Секретариату распространить среди ведомств циркулярное письмо в предложением предоставить таблицы транслитерации и опубликовать полученные материалы в части 7 Справочника ВОИС, как указано в пунктах 20–22 выше.

[Приложение следует]

WIPO STANDARD ST.93

RECOMMENDATIONS ON THE DATA CLEANING OF NAMES

Proposal presented for approval by the Committee on WIPO Standards (CWS)
at its fourteenth session

TABLE OF CONTENTS

INTRODUCTION 2
DEFINITIONS 2
INTAKE 2
TRANSFORMATION OF NAMES 3
VALIDATION AND DISAMBIGUATION 3
MAINTENANCE 4
PUBLICATION AND DATA EXCHANGE 4
STATISTICAL PURPOSES 4
REFERENCES 4

WIPO STANDARD ST.93

RECOMMENDATIONS ON NAME DATA CLEANING

Proposal presented for adoption by the Committee on WIPO Standards (CWS)
at its fourteenth session

Editorial Note by the International Bureau:

Implementation of this Standard is voluntary and non-binding. Each Office may determine, according to its circumstances, whether and when to implement it, either fully or partially, or not at all.

INTRODUCTION

1. This Standard provides general recommendations on the intake, processing, cleaning, and publication of clean name data. This Standard does not provide recommendations on details in relation to approaches to data cleaning, name localization or transformation, such as transliteration, transcription or translation, or approaches to name standardization, such as selection of algorithms, where and when transformations are applied, frequency, or merging strategies. Decisions on such details will vary greatly depending on the party applying the approach, the purpose of transformations, and the quickly evolving nature of matching algorithms.

2. WIPO Standard ST.20 should be referred to for recommendations to produce indexes to patent documents giving names of applicants and other customers, and to promote a uniform presentation of names occurring in name indexes as well as a uniform method of ordering the names in the index itself.

DEFINITIONS

3. In the context of this document:

(a) "IPO" refers to an Intellectual Property Office, which manage the application and registration process for intellectual property rights.

(b) "Customer data" means data on applicants, registrants, owners, legal representatives, or other parties held by an IPO in connection with an IP right, application, registration, or other instrument. This standard is primarily concerned with customer name data: personal names, business names, and related information such as city, address, or email that can be used to disambiguate potential name matches.

(c) "Clean data" means data that is accurate, consistent and reliable. As the degree of cleanness in a large complex data set is difficult to measure, various metrics may be used as proxies for cleanness or related properties, such as fitness for purpose.

(d) "Transliteration" means the mapping of a source language character(s) to a target language (phonetic) character(s).

(e) "Transcription" means the mapping of a source language character/logogram/syllable/phoneme to something that corresponds to the sound in the respective system of the target language.

(f) "Translation" means representing the meaning of a word or concept in the source language with something that corresponds to that meaning in the target language.

(g) "Validation" means the process of conducting automated or manual checks on customer name data to confirm its accuracy, completeness, and compliance.

(h) "Disambiguation" means the processes of comparing names to identify and distinguish identical or similar names that may refer to different entities, so as to eliminate ambiguity in name matching.

INTAKE

4. IPOs may provide the ability for customers to create and manage electronic customer records containing published name information: personal names, business names, names of legal representatives, and related information such as city, address, or email.

5. IPOs should allow a customer record to be associated with multiple applications or registrations for IP rights, so that customers may reuse the same name information for multiple applications or registrations and update their name information in one place.

6. IPOs may provide forms for customers to use to request the IPO to create or change their name or related information. IPOs may allow customers to enter and update their name or related information themselves, or may require a designated party, such as employees, contractors, or an external service to enter and update customer records at the customer's request.

7. Multiple records for one customer may be created and managed by different entities, such as different legal representatives. IPOs should consider this when designing their customer record systems, as multiple records for a single customer may contain slight variations on the same data or be updated at different times by different representatives.

8. IPOs may support entry of the customer's name in native characters of the customer's language, in addition to the customer's name in the language(s) of operation for an IPO, which should be stored using UTF-8 [1]encoding. For instance, an IPO that works in English could allow separate fields for an applicant name in English and the original applicant name in Korean.

9. IPOs may optionally use identification codes to identify customers. Identification codes may be created by the IPO or taken from an external source, such as a registered business number or passport number. Identification codes alone do not resolve issues with clean customer data, such as duplicate entries, name changes, and outdated or incorrect information. IPOs using identification codes should continue to pay attention to and address the considerations in other parts of this Standard.

TRANSFORMATION OF NAMES

10. For data exchange and processing, including the receipt of international applications or registrations, IPOs may consider the name transformation. It is recommended that IPOs should send and receive name data using UTF-8 encoding.

11. It should be noted that the localization or conversion of customer names is extremely prone to error as there are no generally accepted or uniform standards. For localization or transformation of names, there are three ways referred to in this Standard: transliteration, transcription and translation. If IPOs transliterate, transcribe or translate names from one language or character set ~~(~~such as Korean or Latin~~)~~ to another ~~(~~such as English or Cyrillic~~)~~ respectively, they should publish their scheme of transliteration, transcription or translation. If IPOs transform a customer's name, it is recommended to retain the applicant's name in the native characters or language(s), in accordance with paragraph 8 of this Standard. Transliterated, transcribed or translated documents, or parts of documents, should be made available to the customer for review, and customers should have the opportunity to submit corrections if the transliteration, transcription or translation is flawed.

12. Reverse transliteration should be avoided, if possible; instead, it is recommended that the original name should be used instead. For instance, an application filed by "Phony Corp" in Latin characters might be transliterated to Greek characters as "Φονι Κορπ" in an IPO system, and on publication might be reverse transliterated from Greek back to Latin characters as "Foni Corp", leading to mismatches.

VALIDATION AND DISAMBIGUATION

13. Validation and disambiguation approaches should be designed to meet specific objectives, either administrative or statistical, and appropriate methods applied given the objectives. Approaches to name matching and disambiguation should be appropriately scoped and risk assessed in the light of their design objective to ensure appropriate levels of disambiguation are achieved for the use case.

14. IPOs may choose to perform validation of submitted customer information, including automated checks. The validation of such results should be communicated to the customer. Where necessary, any corrections must be approved by the customer prior to implementation in the system. Furthermore, provisions should be made to allow bypassing the automated validation mechanism in instances where it produces inaccurate or incomplete results.

15. IPOs attempting to disambiguate name records (i.e., find duplicate entries) may wish to consider more than just customer names. Names are not inherently unique. For example, there may be multiple individuals named "John Smith" or multiple companies named "Data Corp". Comparing related data points such as city, post code, birthdate, or other information, where available, can increase the likelihood of successful matches.

16. Any validation or disambiguation process initiated by an IPO that could potentially have legal effects, such as correcting or standardizing the name of the registered owner of an IP right, should be confirmed by the customer before the change is made in the IPO's system.

[1] UTF-8 is an encoding system for Unicode.

MAINTENANCE

17. IPOs should develop a strategy to periodically clean data in customer name databases, including searching for and attempting to resolve duplicate records, i.e., multiple records for the same customer. In some instances, duplicates may be merged or combined, for instance, records with slight unintentional differences in spelling such as “ABC Corp” and “ABC Corp.” could be consolidated. In other instances, maintaining separate records might be preferable. Each IPO should decide what approach best fits its own name record management system. The strategy may include involving the customers concerned with the records in the data cleaning process and the responsibility of the cleaned data.

18. IPOs should provide a mechanism for customers to update their name information on multiple applications or IP rights by entering the information once. For instance, this could be achieved by associating each application or IP right with a single customer record containing name information, or by allowing customers to select multiple applications or IP rights and submit one instance of updated name information to be applied to all of them.

19. IPOs may designate someone to be responsible for data cleaning issues, including the development of metrics for measuring data cleanness, regular monitoring and reporting of those metrics, and taking action to improve customer data when needed.

PUBLICATION AND DATA EXCHANGE

20. IPOs should make available updates to name information that are submitted after an IP document has been published. For instance, if “ABC Corp” changes its name to “XYZ Corp” in its customer record, then the name “XYZ Corp” should be associated with the IP right in online publications. The original name may also appear in published IP documents, depending to the legal requirements of the IPO.

21. If an IPO holds other forms of a customer’s name, such as an original name expressed using native characters, these should be included in published IP documents and documents exchanged with other IPOs.

22. If an IPO uses identification codes to identify entities, the codes should be included in published data and data exchanged with other IPOs, unless the identification codes are sensitive and cannot be shared.

STATISTICAL PURPOSES

23. For statistical purposes, IPOs may attempt to match customer data with variations in customer names, or other fields, to achieve counts that are more accurate. In such cases, IPOs should publish their matching strategy or algorithm along with the statistical results so that others can understand the methodology used.

REFERENCES

24. References to the following Standard are of relevance to this Standard:

WIPO Standard ST.20 Recommendation for the preparation of name indexes to patent documents

[End of the proposed Standard]

[Конец приложения и документа]