

## ДОДАТОК II

### ВИЗНАЧЕННЯ ТИПУ ДОКУМЕНТА (DTD) ДЛЯ ПЕРЕЛІКУ ПОСЛІДОВНОСТЕЙ

#### Редакція 1.3

*Редакція, схвалена Комітетом зі стандартів ВОІВ (КСВ)  
на його восьмій сесії 4 грудня 2020 року*

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<!--Додаток II до стандарту ВОІВ ST.26, визначення типу документа (DTD) для переліку  
послідовностей
```

Цю сутність можна ідентифікувати за допомогою кваліфікатора "PUBLIC":

```
*****  
PUBLIC "-//WIPO//DTD SEQUENCE LISTING 1.3//EN" "ST26SequenceListing_V1_3.dtd"  
*****
```

```
*****
```

\* URL-адреса публічного DTD

\* [http://www.wipo.int/standards/DTD/ST26SequenceListing\\_V1\\_2.dtd](http://www.wipo.int/standards/DTD/ST26SequenceListing_V1_2.dtd)

```
*****
```

\*Перегляд Додатку II до стандарту ВОІВ ST.26 був затверджений Комітетом зі стандартів  
ВОІВ (КСВ) на його шостій сесії.

```
*****
```

\* КОНТАКТНА ІНФОРМАЦІЯ

```
*****
```

\*

\*xml.standards@wipo.int

\*

```
*****
```

\* ПРИМІТКИ

```
*****
```

\*Частина з даними щодо послідовності є підрозділом INSDC DTD V.1.5, що містить лише  
вимоги за стандартом

ВОІВ ST.26.\*\*\*\*\*

\*\*

\* ІСТОРІЯ ПЕРЕГЛЯДУ

```
*****
```

20-05-20: Редакція 13 затверджена на восьмій сесії КСВ.

Зміни:

- Необов'язковий атрибут originalFreeTextLanguageCode, доданий до  
<ST26SequenceListing>, щоб дозволити заявникам вказати мову довільного тексту в  
оригінальному переліку послідовностей.
- Необов'язковий атрибут originalFreeTextLanguageCode, доданий до  
<ST26SequenceListing>, щоб дозволити заявникам вказувати мову довільного тексту в  
елементі <NonEnglishQualifier\_value>.
- Необов'язковий атрибут id, доданий до INSDQualifier для полегшення порівняння  
значень, що залежать від мови, між переліками послідовностей.
- Необов'язковий елемент <NonEnglishQualifier\_value>, доданий до елемента  
<INSDQualifier>, щоб дозволити заявникам набирати неанглійською мовою  
кваліфікатори, що залежать від мови, за допомогою символів, зазначених у параграфі  
40(а) основного тексту стандарту ST.26.

2018-10-19: Редакція 1.2 затверджена на шостій сесії КСВ.

Зміни:

<INSDQualifier\*> замінено на <INSDQualifier+> для приведення у відповідність з бізнес-потребами та рекомендаціями Національного центру біотехнологічної інформації (NCBI) (елемент INSDFeature\_qual (за наявності) повинен мати один або більше елементів INSDQualifier).

\*\*\*\*\*

2017-06-02: Редакція 1.1 затверджена на п'ятій сесії КСВ

Зміни:

Додано примітки до <INSDSeq\_length>, <INSDSeq\_division> and <INSDSeq\_sequence> для пояснення причини відмінностей між INSDC DTD v.1.5 та Переліком послідовностей за стандартом BOIB ST.26 DTD V1\_1.

\*\*\*\*\*

2016-03-24: Редакція 1.0 затверджена на відновленій четвертій сесії КСВ

2014-03-11: Остаточний проект для затвердження.

\*\*\*\*\*

ST26SequenceListing

\*\*\*\*\*

\* КОРЕНЕВИЙ ЕЛЕМЕНТ

\*\*\*\*\*

-->

```
<!ELEMENT ST26SequenceListing ((ApplicantFileReference | (
    ApplicationIdentification,ApplicantFileReference?)),
    EarliestPriorityApplicationIdentification?, (ApplicantName,
    ApplicantNameLatin?)), (InventorName,InventorNameLatin?)),
    InventionTitle+,SequenceTotalQuantity,SequenceData+)>
```

<!--Елементи "ApplicantName" та "InventorName" є необов'язковими у цьому DTD для спрощення конвертації між різними схемами кодування-->

<!--originalFreeTextLanguageCode:

Код мови (див. посилання в параграфі 9 на ISO 639-1: 2002) для однієї оригінальної мови, на якій кваліфікатори довільного тексту, залежні від мови (NonEnglishQualifier\_value)

були підготовлені.

-->

<!-- nonEnglishFreeTextLanguageCode:

Код мови (див. посилання в параграфі 9 на ISO 639-1: 2002) для мови, на якій наразі, залежні від мови, відповідають кваліфікатори довільного тексту (NonEnglishQualifier\_value).

-->

```
<!ATTLIST ST26SequenceListing
    dtdVersion CDATA #REQUIRED
    fileName CDATA #IMPLIED
    softwareName CDATA #IMPLIED
    softwareVersion CDATA #IMPLIED
    productionDate CDATA #IMPLIED
    origianlFreeTextLanguageCode CDATA #IMPLIED
    nonEnglishFreeTextLanguageCode CDATA #IMPLIED
```

>

<!--ApplicantFileReference

Зазначення заявника або представника; обов'язковий за відсутності ідентифікації заявки.

-->

```
<!ELEMENT ApplicantFileReference (#PCDATA) >
```

<!--ApplicationIdentification

Ідентифікація заявки, стосовно якої подано перелік послідовностей.

-->

```
<!ELEMENT ApplicationIdentification (IPOfficeCode,ApplicationNumberText,
    FilingDate?) >
```

<!--EarliestPriorityApplicationIdentification

Ідентифікація заявки з більш раннім пріоритетом, що містить елементи "IPOfficeCode", "ApplicationNumberText" і "FilingDate".

-->

```
<!ELEMENT EarliestPriorityApplicationIdentification (IPOfficeCode,
ApplicationNumberText,FilingDate?)>

<!--ApplicantName
Ім'я першого вказаного заявника, представлене символами, визначеними у параграфі 40(a)
основної частини стандарту BOIB ST.26.
-->
<!--languageCode: Відповідний код мови за стандартом ISO 639-1 "Коди для представлення
назв мов. Частина 1: Код Альфа-2"
-->
<!ELEMENT ApplicantName (#PCDATA)>
<!ATTLIST ApplicantName
languageCode CDATA #REQUIRED
>

<!--ApplicantNameLatin
Якщо елемент "ApplicantName" наведено символами, не зазначеними у параграфі 40(b),
необхідно навести переклад або транслітерацію імені першого зазначеного заявника
символами, зазначеними у параграфі 40(b) основної частини стандарту BOIB ST.26.
-->
<!ELEMENT ApplicantNameLatin (#PCDATA)>

<!--InventorName
Ім'я першого вказаного винахідника, представлене символами, визначеними у
параграфі 40(a).
-->
<!--languageCode: Відповідний код мови за стандартом ISO 639-1 "Коди для представлення
назв мов. Частина 1: Код Альфа-2"
-->
<!ELEMENT InventorName (#PCDATA)>
<!ATTLIST InventorName
languageCode CDATA #REQUIRED
>

<!--InventorNameLatin
Якщо елемент "InventorName" наведено символами, не зазначеними у параграфі 40(b),
необхідно навести переклад або транслітерацію імені першого зазначеного винахідника
символами, зазначеними у параграфі 40(b).
-->
<!ELEMENT InventorNameLatin (#PCDATA)>

<!--InventionTitle
Назва винаходу, вказана символами як зазначено у параграфі 40(a) основної частини
стандарту ST.26, мовою оригіналу. Переклад назви винаходу додатковими мовами може
наводитися символами як зазначено у параграфі 40(a) з використанням додаткових елементів
"InventionTitle". Важано не більше 2-7 слів.
-->
<!--languageCode: Відповідний код мови за стандартом ISO 639-1 "Коди для представлення
назв мов. Частина 1: Код Альфа-2"
-->
<!ELEMENT InventionTitle (#PCDATA) >
<!ATTLIST InventionTitle
languageCode CDATA #REQUIRED
>

<!--SequenceTotalQuantity
Вказує на загальну кількість послідовностей у документі.
Має бути легкодоступним для автоматичної обробки.
-->
<!ELEMENT SequenceTotalQuantity (#PCDATA)>

<!--SequenceData
Дані стосовно певної послідовності.
Інформація стосовно навмисно пропущених послідовностей міститься в основній частині
стандарту BOIB ST.26.
-->
```

```

<!ELEMENT SequenceData (INSDSeq)>
<!--ATTLIST SequenceData
sequenceIDNumber CDATA #REQUIRED
-->

<!--IPOfficeCode
Код за стандартом BOIB ST.3. Наприклад, якщо ідентифікацією заявки є "PCT/IB2013/099999",
значенням "IPOfficeCode" є "MB" для позначення Міжнародного бюро BOIB.
-->
<!--ELEMENT IPOfficeCode (#PCDATA)>

<!--ApplicationNumberText
Ідентифікація заявки, надана відомством подання (наприклад, PCT/IB2013/999999)
-->
<!--ELEMENT ApplicationNumberText (#PCDATA)>

<!--FilingDate
Дата подання патентної заявки, стосовно якої подається перелік послідовностей, у форматі
"PPPP-MM-DD" за стандартом BOIB ST.2, де "PPPP" - календарний рік, MM - календарний
місяць, DD - день у межах календарного місяця, наприклад 2015-01-31. Детальну інформацію
викладено у параграфах 7(a) і 11 стандарту BOIB ST.2.
-->
<!--ELEMENT FilingDate (#PCDATA)>

<!--
*****
* Частина INSD
*****

Метою частини INSD цього DTD є визначення специфічних DTD для переліків послідовностей
для підтримки роботи відомств промислової власності та спрощення обміну даними з
публічними депозитаріями.

Частина INSD є підмножиною INSD DTD v1.5 і може використовуватись лише для генерування
екземпляра XML, оскільки вона не підтримує всю структуру INSD.

Ця частина заснована на:

Проекті міжнародного співробітництва баз даних нуклеотидних послідовностей (INSDC).

"INSDSeq" містить елементи послідовності у форматах, як їх зазначено в GenBank/EMBL/DDBJ.
Тут використовуються не всі елементи.
-->

<!--INSDSeq
Дані щодо послідовності. Змінено елементи "INSD V1.5 DTD", "INSDSeq_division" та
"INSDSeq_sequence" з необов'язкових на обов'язкові для бізнес-потреб.
-->
<!--ELEMENT INSDSeq (INSDSeq_length,INSDSeq_moltype,INSDSeq_division,
INSDSeq_other-seqids?,INSDSeq_feature-table?,INSDSeq_sequence)>

<!--INSDSeq_length
Довжина послідовності. "INSDSeq_length" дозволяються тільки цілі числа.
-->
<!--ELEMENT INSDSeq_length (#PCDATA)>

<!--INSDSeq_moltype
Допустимі значення: DNA, RNA, AA
-->
<!--ELEMENT INSDSeq_moltype (#PCDATA)>

<!--INSDSeq_division
Зазначення того, що послідовність відноситься до патентної заявки. Повинен бути
заповнений значенням "PAT".
-->

```

---

```
<!ELEMENT INSDSeq_division (#PCDATA)>
```

```
<!--INSDSeq_other-seqids
```

В рамках обміну даними з провайдером баз даних, патентні відомства мають заповнювати елемент "INSDSeq\_other-seqids" для кожної послідовності одним елементом "INSDSeqid" з посиланням на відповідний опублікований патент та ідентифікаційними даними послідовності.

```
-->
<!ELEMENT INSDSeq_other-seqids (INSDSeqid?)>
```

```
<!--INSDSeq_feature-table
```

Інформація про локалізацію та призначення різних ділянок у межах певної послідовності. Елемент "INSDSeq\_feature-table" повинен завжди містити принаймні одну характеристику.

```
-->
<!ELEMENT INSDSeq_feature-table (INSDFeature+)>
```

```
<!--INSDSeq_sequence
```

Залишки послідовності. Послідовність не повинна містити цифри, розділові знаки або пробіли.

```
-->
<!ELEMENT INSDSeq_sequence (#PCDATA)>
```

```
<!--INSDSeqid
```

Призначений для використання патентними відомствами лише під час обміну даними.

Формат:

```
pat|{код відомства}|{номер публікації}|{код виду документа}|{ідентифікаційний номер послідовності}
```

де кодом відомства є код відомства промислової власності, яке опублікувало патентний документ; номером публікації є номер опублікованої заявки або патенту; кодом виду документа є буквенний код для ідентифікації патентного документа відповідно до стандарту BOIB ST.16; а ідентифікаційним номером послідовності є номер послідовності в заявці або у патенті

Приклад:

```
pat|WO|2013999999|A1|123456
```

Це 123456-а послідовність у патентній публікації WO2013999999 (A1)

```
-->
<!ELEMENT INSDSeqid (#PCDATA)>
```

```
<!--INSDFeature
```

Опис однієї функції.

```
-->
<!ELEMENT INSDFeature (INSDFeature_key,INSDFeature_location,INSDFeature_qualifiers?)>
```

```
<!--INSDFeature_key
```

Слово чи абревіатура, що ідентифікує функцію.

```
-->
<!ELEMENT INSDFeature_key (#PCDATA)>
```

```
<!--INSDFeature_location
```

Ділянка представленої послідовності, що відповідає функції.

```
-->
<!ELEMENT INSDFeature_location (#PCDATA)>
```

```
<!--INSDFeature_qualifiers
```

Перелік кваліфікаторів з додатковою інформацією про функцію.

```
-->
<!ELEMENT INSDFeature_qualifiers (INSDQualifier*)>
```

```
<!--INSDQualifier
```

Додаткова інформація про функцію.

Інформація про кодуючі послідовності та варіанти міститься в основній частині стандарту BOIB ST.26.

```

-->
<!--id
Унікальний ідентифікатор INSDQualifier для полегшення порівняння версій списків
последовностей, особливо залежних від мови значень кваліфікаторів на різних мовах.
-->
<!--ELEMENT INSDQualifier (INSDQualifier_name, INSDQualifier_value?,
NonEnglishQualifier_value?)>
<!--ATTLIST INSDQualifier
            id ID #IMPLIED
>
<!--INSDQualifier_name
Ім'я кваліфікатора.
-->
<!--ELEMENT INSDQualifier_name (#PCDATA)>
<!--INSDQualifier_value
Значення кваліфікатора. Якщо кваліфікатор залежить від мови, його значення повинно бути
англійською мовою та введено символами, зазначених в пункті 40 (b).
→
<!--ELEMENT INSDQualifier value (#PCDATA)>
<!--NonEnglishQualifier_value
Значення залежного від мови кваліфікатора в мові, яка не є англійською мовою та введена
за допомогою
символів, зазначених в пункті 40 (a). Мова позначається атрибутом
nonEnglishFreeTextLanguageCode.

-->
<!--ELEMENT NonEnglishQualifier_value (#PCDATA)>

```

[Додаток III наведено далі]