

## **Технические требования для подключения к «Национальному агрегатору открытых репозиториях российских университетов» (НОРА)**

НОРА представляет из себя архивное хранилище открытых данных. Для депонирования своих данных университетское хранилище должно соответствовать ряду требований:

1. Данные должны быть представлены в унифицированном виде: полные тексты и метаданные.
2. Метаданные должны содержать, как минимум, следующие поля:
  - название документа;
  - URL - прямая ссылка на скачивание полного текста;
  - дату создания документа;
  - инициалы и фамилии всех авторов (либо явное указание на отсутствие авторов).
3. Любые другие поля опциональны, однако рекомендуется использовать поля из базовой схемы Dublin Core.
4. Репозиторий должен реализовывать на своей стороне OAI-PMH сервер для передачи метаданных агрегатору (<https://www.openarchives.org/pmh/>).
5. Полные тексты должны быть доступны для скачивания по ссылке из метаданных. Также должна быть отключена защита от массового скачивания для IP-адреса агрегатора.

Ниже будет описано возможное решение для построения репозитория, готового к подключению к НОРА, с использованием ПО с открытым исходным кодом DSpace.

DSpace является программным обеспечением для создания институциональных репозиториями. На момент написания настоящей документации актуальной является версия DSpace 6.2. Для установки DSpace необходим сервер с минимум 8 ГБ оперативной памяти, процессор 2.2 ГГц, 2 ГБ свободного места для установки программы и дополнительное место для хранения загружаемых данных. Процесс установки DSpace подробно описан в инструкции:

<https://wiki.duraspace.org/display/DSDOC6x/Installing+DSpace>.

По умолчанию ссылка на полный текст не включена во встроенную схему. Для её включения необходимо добавить в файл {директория установки dspace}/config/crosswalks/oai/metadataFormats/oai\_dc.xsl следующие строки в конце секции oai\_dc:dc (начиная с 4 строки с конца):

```
<!-- dc.url -->
<xsl:for-each
select="doc:metadata/doc:element[@name='bundles']/doc:element[@name='bundle']">
  <xsl:if test="doc:field[@name='name']/text() = 'ORIGINAL'">
    <xsl:for-each
select="doc:element[@name='bitstreams']/doc:element[@name='bitstream']">
      <dc:url>
        <xsl:attribute name="type">
          <xsl:value-of
select="doc:field[@name='format']/text()"></xsl:value-of>
        </xsl:attribute>
        <xsl:attribute name="length">
          <xsl:value-of
select="doc:field[@name='size']/text()"></xsl:value-of>
        </xsl:attribute>
```

```
of>
    <xsl:value-of select="doc:field[@name='url']/text()"></xsl:value-
of>
    </dc:url>
  </xsl:for-each>
</xsl:if>
</xsl:for-each>
```

После чего необходимо перезапустить контейнер с DSpace и инвалидировать кэш OAI:  
dspace oai clean-cache

После установки DSpace будет готов к использованию как хранилище данных с возможностью поиска, добавления новых данных, а также OAI-сервер для предоставления данных внешним сервисам, в том числе и харвестеру HORA. Процесс запуска OAI-сервера описан в этой инструкции: <https://wiki.duraspace.org/display/DSDOC6x/OAI+2.0+Server>

Для подключения к HORA необходимо сообщить адрес вашего OAI-сервера системному администратору. Для самостоятельной проверки корректной настройки OAI рекомендуется использовать сервис валидации OAI: <http://validator.oaipmh.com/>

После установки DSpace и подключению к HORA все данные, загружаемые в локальные репозиторий, будут автоматически синхронизироваться с HORA.