

# Корпус публикаций журнала «Современник» (1847–1866)

Е. И. Вожик

## Аннотация

В корпус входят тексты публикаций журнала «Современник» (1847–1866); каждая публикация снабжена метаданными, основной набор которых воспроизводится по указателю содержания журнала, составленному В. Э. Боградом (*Боград В. Э. Журнал «Современник». 1847–1866: Указатель содержания. М.; Л.: ГИХЛ, 1959*); сведения указателя проверены, уточнены, реструктурированы и дополнены по оригинальным выпускам «Современника». Тексты публикаций объединены в zip-архиве `publications_txt`, где каждая публикация представлена в виде отдельного `.txt` файла (кодировка UTF-8); имя файла содержит указание на год издания, номер журнального выпуска и порядковый номер публикации в выпуске. Метаданные к текстам публикаций находятся в таблице `publications_metadata.tsv`; метаданные к журнальным приложениям представлены в таблице `supplement_metadata.tsv`. Текстовые файлы с публикациями соотносятся с таблицей `publications_metadata.tsv` согласно сведениям, содержащимся в столбце `filename_txt`. Принципы представления текстов в корпусе, а также подробное описание структуры таблиц метаданных приведены в файле `readme.md`.

Формат цитирования датасета: *Вожик Е. И. Корпус публикаций журнала «Современник» (1847–1866) // Репозиторий открытых данных по русской литературе и фольклору. 2023. V1. DOI: <https://doi.org/10.31860/openlit-2023.11-C006>*

## Общее описание данных

В корпус входят тексты публикаций журнала «Современник» (1847–1866). Публикацией здесь и далее называется любое художественное, научное, критическое, публицистическое или иное произведение, напечатанное на страницах журнала; содержательные, формальные и любые другие особенности произведений, которые вошли в состав «Современника», при выборе родового обозначения *публикация* не учитывались.

Каждая публикация снабжена метаданными, основной набор которых воспроизводится по указателю содержания журнала, составленному В. Э. Боградом (*Боград В. Э. Журнал «Современник». 1847–1866: Указатель содержания*

ния. М.; Л.: ГИХЛ, 1959); сведения указателя проверены, уточнены, реструктурированы и дополнены по оригинальным выпускам «Современника».

В ходе проверки сведений, содержащихся в указателе Богграда, и заполнения таблиц метаданных к публикациям журнала и журнальным приложениям (`publications_metadata.tsv` и `supplement_metadata.tsv`) были просмотрены все номера «Современника», выходившие с 1847 по 1866 гг., в том числе последний, пятый выпуск журнала за 1866 г., часть тиража которого была изъята из типографии и уничтожена по личному распоряжению Александра II.

В то время как главной целью указателя к «Современнику» 1959 г. было «максимально возможное раскрытие авторов анонимных и псевдонимных произведений» (Боград В. Э. Указ. соч. С. 5), составительница настоящего корпуса руководствовалась стремлением репрезентировать принятый в «Современнике» способ представления данных разного типа и, в частности, установить особенности фигурирования в журнале авторского имени.

Так, Боград последовательно исключал из заголовков публикаций входящие в них имена сочинителей. Напр., вместо «Рассказы Жерара о львиной охоте» (Современник. 1852. № 2. Смесь. С. 251–265) в указателе значится: «Рассказы о львиной охоте». В таблицах метаданных к корпусу заголовки публикаций, равно как и подписи к ним, восстановлены в соответствии с данными журнала.

Особое внимание при подготовке корпуса было уделено вопросу разделения журнальных материалов на отдельные публикации. Ряд решений, принятых Боградом в отношении границ тех или иных публикаций, был пересмотрен. Наиболее трудным оказалось определение статуса стихотворений одного и того же автора, публикуемых совместно.

Если стихотворения 1) были изданы под общим заголовком, в том числе обозначенным в оглавлении (напр., «Пять стихотворений графа А. К. Толстого» или «Два последние стихотворения Жуковского»), и/или 2) имели одну общую подпись (см. стихотворения Н. Ф. Щербины «Музам» и «Notturmo» в 3 номере «Современника» за 1851 г.), такой стихотворный цикл признавался одной статьей.

Напротив, указанием на самостоятельность стихотворных произведений, следующих в журнале друг за другом, считалось 1) отсутствие у них общего заголовка, а также 2) наличие отдельной подписи под каждым из текстов. Таков, напр., случай стихотворной подборки Н. А. Некрасова, опубликованной в 9 номере «Современника» за 1856 г.: под каждым из трех стихотворений («Я посетил твоё кладбище...», «Самодовольных болтунов...», «Давно, отвергнутый тобою...») значится здесь имя сочинителя.

Сведения о произведениях, являющихся приложениями к «Современнику», представлены в таблице `supplement_metadata.tsv`. Приложением считается произведение (или его часть), обозначенное в качестве приложения

в оглавлении к журнальному выпуску, четко выделяемое из состава журнала по формальным признакам (отдельная пагинация, увеличенный шрифт, отсутствие колонтитулов) и нередко имеющее собственный титульный лист.

В число приложений не входят «Литературный ералаш», «Свисток», «Устройство быта помещичьих крестьян», осмысленные и представленные в таблице `publications_metadata.tsv` как отделы журнала, близкие по значению отделам «Словесность», «Науки и художества», «Критика и библиография» и т. д. В пользу такого решения свидетельствует структура оглавлений к журнальным выпускам.

## **Полнота данных**

### **Метаданные**

В таблице метаданных (`publications_metadata.tsv`) к текстам журнальных статей исправлены неточности, выявленные в издании Бограда. Например, учтена публикация «Из наблюдений наставника семинарии» (Современник. 1862. № 3. Словесность, науки и художества. С. 277–288), отсутствующая в указателе к журналу, или восстановлен подзаголовок статьи «Письма об Осташкове. Образец городского устройства в России» (Современник. 1862. № 5. Словесность, науки и художества. С. 49–86).

Купюры, используемые в указателе Бограда, раскрываются; исключение составляют особо длинные названия публикаций, состоящие более чем из 10 строк. Сокращения, которые встречаются в обеих таблицах метаданных (`publications_metadata.tsv` и `supplement_metadata.tsv`), воспроизводятся строго по тексту журнала; унифицированно отображаются сокращения от названий городов (М.; СПб.).

Объявления об издании «Современника», повторяющиеся в нескольких номерах журнала, приводятся в таблице `publications_metadata.tsv` только один раз, по месту первой публикации. Прочие объявления в таблице метаданных не отражены.

Все записи в таблицах метаданных приведены в соответствие с нормами современной орфографии и пунктуации.

### **Тексты**

С 1847 по 1866 гг. в журнале «Современник» было опубликовано 5412 материалов (без учета приложений к журналу); в корпусе представлены тексты для 4686 из них. Тексты приложений в корпус не включены.

Тексты, составившие содержание корпуса, были распознаны с помощью OCR и извлечены из .pdf версий журнальных выпусков и томов журнала, хранящихся в электронной библиотеке Google Books. Из полученных

текстовых файлов были автоматически удалены разделители страниц, а также пустые строки; особенности дореволюционной орфографии были сохранены.

Тексты «Современника» не вычитывались. В силу того что качество распознавания отдельных выпусков журнала оказалось невысоким, было принято решение не исключать из состава корпуса верхние и нижние колон-тителы, номера страниц, а также паратекстовые элементы журнальных публикаций (заголовки, подписи, даты окончания работы над статьей и т. д.).

В текстовых файлах корпуса отсутствуют:

1. титульные листы и обороты титула,
2. сведения о допуске журнального выпуска к печати, даты цензурного разрешения, имена цензоров,
3. рекламные объявления, а также объявления о выходе новых книг и о начале, возобновлении, прекращении периодических изданий.

## **Формат и структура данных**

Данные представлены в формате `tab-separated values` в файлах `publications_metadata.tsv` и `supplement_metadata.tsv` и `zip`-архиве `publications_txt`.

Для таблицы `publications_metadata.tsv` единицей данных (соответствующей строке в таблице) является журнальная публикация; для таблицы `supplement_metadata.tsv` — приложение к выпуску журнала, представляющее собой самостоятельное произведение или его часть.

Тексты журнальных публикаций в виде отдельных `.txt` файлов (кодировка UTF-8) объединены в `zip`-архиве `publications_txt`; их названия содержат три разделенных символом подчеркивания числа (или числовых диапазона) (`1847_1_1.txt`, `1863_1-2_1.txt` и т. д.), соответствующих году издания журнала, номеру журнального выпуска и порядковому номеру публикации в выпуске соответственно. Названия `.txt` файлов приведены в столбце `filename_txt` таблицы `publications_metadata.tsv`.

Данные проверены на целостность автоматизированным тестированием. Тесты и их описание приведены в подкаталоге `tests`.

### **Структура таблицы `publications_metadata.tsv`**

- `id` — порядковый номер публикации в таблице метаданных.
- `uid` — идентификатор публикации (для публикаций, не образующих цикла) или цикла публикаций. Общность `uid`'а позволяет установить принадлежность циклу произведений, имеющих разные заголовки:

например, цикл образуют очерки «Женихи бурсы», «Бегуны и спасенные бурсы» и «Переходное время бурсы» Н. Г. Помяловского.

- **author** — фамилия и инициалы имени и отчества автора(-ов) публикации согласно сведениям указателя Богграда. Если авторство публикации составителем указателя не установлено, в столбце значится слово *неизвестный*. Имена авторов представлены единообразно, в формате *Тургенев И. С.*
- **attribution** — дополнительные сведения об атрибуции публикации согласно данным указателя Богграда: ее авторство может быть 1) установлено, 2) установлено частично (для статей, написанных в соавторстве), 3) не установлено, 4) под вопросом (для случаев, когда авторство устанавливается предположительно).
- **signature** — имя или псевдоним автора, составителя или переводчика публикации, располагающиеся по ее завершении. Наличие или отсутствие подписи определяется строго формально: указания на имя автора, содержащиеся в подзаголовке или примечаниях к публикации, в этом столбце не отражены. Отсутствие подписи к публикации обозначается словом *нет*.
- **signature\_index** — имя или псевдоним автора, составителя или переводчика публикации, указанные в оглавлении. В столбце приводятся подписи, явно выделяющиеся из заголовка публикации (курсивом, размером шрифта). Отсутствие подписи к публикации обозначается словом *нет*.
- **title** — полное заглавие публикации, включающее также подзаголовок и имя автора, не отделимое от него (при наличии). Заглавие указывается по тексту публикации, в большинстве случаев превосходящему по объему вариант заглавия в тексте оглавления. Если публикация не имеет заглавия, оно воспроизводится условно, по ее содержанию, и заключается в треугольные скобки. В скобки угловой формы также помещаются раскрываемые Богградом псевдонимы и криптонимы. Кроме того, в столбце отражены названия подразделов внутри той или иной публикации (см. серию статей «Литературные мелочи» Постороннего сатирика), а также заголовки отдельных стихотворений, входящих в состав цикла.
- **part** — унифицированное числовое обозначение, заменяющее словесное указание на порядковый номер публикации в цикле или серии («статья первая», «письмо пятое и последнее» и т. п.). В некоторых случаях вхождение тех или иных публикаций в один цикл устанавливается на основании общности их заголовка или эксплицитного указания на это вхождение, содержащегося в тексте публикации (см. «Жизнь Алкивиада, по новейшим о ней исследованиям» В. А. Милютина или «Антропологический принцип в философии» Н. Г. Чернышевского). Фельетоны, регулярно печатающиеся в журнале под

общим заглавием, но не имеющие порядковых номеров (см. фельетоны о Новом поэте, написанные И. И. Панаевым — единолично или в соавторстве, — или заметки из цикла «Наша общественная жизнь» М. Е. Салтыкова-Щедрина), многочастными публикациями не признаются: по значению такие фельетоны близки журнальной рубрике, материал которой не имеет строгих пределов и потенциально не может быть исчерпан. Если публикация не является составной частью цикла, в столбце значится слово *нет*.

- *translator* — содержащееся в журнале указание на имя переводчика публикации. Если таковое указание в тексте журнального выпуска отсутствует, в столбце ставится слово *нет*. Данные этого столбца не позволяют установить полный состав оригинальных, переводных или компилятивных произведений, публиковавшихся в «Современнике»; в ряде случаев провести это различие оказывается принципиально невозможным в силу особенностей обращения с чужим словом, принятого среди литераторов и журналистов эпохи. Написание имен переводчиков унифицировано и приведено к формату *Кронеберг А. И.*
- *dedication* — посвящение, значащееся в подзаголовке публикации (при наличии). В столбце отмечены также шуточные, пародийные посвящения (см. «Requiem» Г. З. Елисеева, имеющий подзаголовок «Рано погибшим надеждам»). Если публикация не имеет посвящения, в столбце значится слово *нет*.
- *type* — тип речевой организации текста публикации (проза или стихи). Прозаические публикации, содержащие стихотворный текст (см. серию статей «Русские второстепенные поэты»), отнесены к типу *проза*.
- *year* — год выхода номера и тома журнала, в состав которых входит публикация.
- *publication* — дата фактического выхода в свет журнального номера, установленная Бюро на основании объявлений, которые публиковались в петербургских газетах. Если дата не установлена точно, в столбце значится *неизвестно*; для пятого номера «Современника» за 1866 г., не допущенного к распространению, в столбце стоит слово *нет*.
- *sensorship* — дата цензурного разрешения, помещаемая на обороте титульного листа (для нечетных номеров) или в оглавлении тома журнала (для четных номеров), а с конца 1855 г. — дополнительно на с. 3 обложки. Если в журнале указаны две даты цензурного разрешения, они приводятся в столбце через запятую. Если дата цензурного разрешения в журнале не указана, в столбце значится *не указано*.
- *sensor* — фамилия, инициалы имени и отчества цензора(-ов), пропустившего(-их) номер журнала в печать. Если имя цензора в

журнале не указано, в столбце значится *не указано*. Имена цензоров представлены единообразно, в формате *Крылов А. Л.*

- `volume` — номер тома журнала, в состав которого входит публикация; обозначается римскими цифрами.
- `number` — номер журнального выпуска, в состав которого входит публикация; обозначается арабскими цифрами.
- `section_content` — название раздела журнала, в который входит публикация, восстанавливаемое по ее содержанию в соответствии с традицией наименования журнальных разделов (при отсутствии эксплицитного указания на вхождение в тот или иной раздел). Публикации, содержащие сведения об издании «Современника», некрологи сотрудников журнала, официальные распоряжения и т. п. условно отнесены к разделу «От редакции»; публикации, освещавшие хронику военных событий 1855–1856 гг. и имеющие отдельную пагинацию, объединены в раздел «Хроника военных событий». Перечень допустимых значений:
  - Библиография
  - Иностранная литература
  - Критика
  - Критика и библиография
  - Литературный ералаш
  - Моды
  - Науки и искусства
  - От редакции
  - Свисток. Собрание литературных, журнальных и других заметок
  - Словесность
  - Словесность, науки и искусства
  - Смесь
  - Современное обозрение
  - Современные заметки
  - Устройство быта помещичьих крестьян
  - Хроника военных действий.
- `section_form` — название раздела журнала, в который входит публикация, указанное в журнальном выпуске (в верхних колонтитулах или оглавлении). Если название раздела в журнале не указано, в столбце значится *не указано*. Перечень допустимых значений:
  - Библиография
  - Иностранная литература
  - Критика
  - Критика и библиография
  - Литературный ералаш
  - Моды
  - Науки и искусства

- Свисток. Собрание литературных, журнальных и других заметок
  - Словесность
  - Словесность, науки и художества
  - Смесь
  - Современное обозрение
  - Современные заметки
  - Устройство быта помещичьих крестьян
  - не указано.
- `first_page` — первая страница публикации.
  - `last_page` — последняя страница публикации.
  - `length` — длина публикации в страницах.
  - `filename_txt` — название .txt файла, содержащего текст публикации.
  - `comments` — комментарии, представляющие собой типологически разнородные метаданные к публикациям журнала (сведения об альтернативных — проставленных на части тиража — датах цензурного разрешения, редакторах, внутренних подзаголовках публикаций, нарушениях пагинации и т. д.).

#### Структура таблицы `supplement_metadata.tsv`

- `id` — порядковый номер публикации-приложения в таблице метаданных.
- `uid` — идентификатор публикации-приложения (для приложений, не образующих цикла) или части произведения, являющегося приложением к журналу. Идентификатор состоит из литеры S (supplement) и порядкового номера приложения в таблице.
- `author` — фамилия и инициалы имени и отчества автора(-ов) публикации-приложения согласно сведениям указателя Бограда. Если авторство публикации составителем указателя не установлено, в колонке ставится слово *неизвестный*. Имена авторов представлены в этом столбце единообразно, в формате *Диккенс Ч.*
- `title` — полное заглавие публикации-приложения, включающее также подзаголовки и имя автора, которое является его неотделимой частью. Как правило, заголовки приложений в тексте выпуска и в оглавлении к нему приводятся в журнале единообразно.
- `part` — число, заменяющее словесное обозначение порядкового номера приложения в серии публикаций («часть первая», «очерк второй» и т. п.). Если публикация не является составной частью цикла, в столбце значится слово *нет*.



- `translator` — содержащееся в журнале указание на имя переводчика произведения, являющегося приложением к журналу. Если таковое указание в тексте журнального выпуска отсутствует, в столбце ставится слово *нет*.
- `year` — год публикации журнального выпуска, в состав которого входит приложение.
- `number` — номер журнального выпуска, в составе которого публикуется приложение; обозначается арабскими цифрами.

### **Благодарности**

Составительница корпуса благодарит Б. В. Орехова и Р. А. Лисюкова за помощь в подготовке текстов журнальных выпусков, а также выражает признательность сотрудникам библиотеки ИРЛИ РАН, оказавшим большое содействие работе с оригинальными экземплярами «Современника».