

Министерство науки и высшего образования Российской Федерации
Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и технологий

Работа допущена к защите
Зам. директора ВШИСиСТ ИКНТ
_____ А. В. Щукин
«_____» _____ 2022 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ
АВТОМАТИЧЕСКАЯ ОБРАБОТКА И ГЕНЕРАЦИЯ ТЕКСТА НА
ЕСТЕСТВЕННЫХ ЯЗЫКАХ С ПРИМЕНЕНИЕМ ИСКУССТВЕННЫХ
НЕЙРОННЫХ СЕТЕЙ**

по направлению подготовки 02.04.03.01 Математическое обеспечение и администрирование корпоративных информационных систем

Направленность (профиль) 02.04.03.01_YY Наименование направленности (профиля) образовательной программы

Выполнил
студент гр. 3540203/00101

В. К. Фурман

Руководитель
доцент каф. ВШИСиСТ,
к. ф.-м. н.

В. Г. Пак

Консультант
по нормоконтролю

В. А. Пархоменко

Санкт-Петербург
2022

СОДЕРЖАНИЕ

Введение	3
Глава 1. Название первой главы: всестороннее изучение объекта и предмета исследования, анализ результатов, полученных другими авторами	4
1.1. Название параграфа	4
1.1.1. Название первого подпараграфа первого параграфа первой главы для демонстрации переноса слов в содержании	4
1.2. Название параграфа	4
1.3. Выводы	6
Глава 2. Название второй главы: разработка метода, алгоритма, модели исследования	6
2.1. Название параграфа	6
2.2. Название параграфа	7
2.2.1. Название подпараграфа	7
2.3. Название параграфа	8
2.4. Выводы	15
Глава 3. Название третьей главы: разработка программного обеспечения .	16
3.1. Название параграфа	16
3.2. Название параграфа	16
3.3. Выводы	16
Глава 4. Название четвёртой главы. Апробация результатов исследования, а именно: метода, алгоритма, модели исследования	16
4.1. Название параграфа	17
4.2. Название параграфа	17
4.3. Выводы	17
Заключение	18
Словарь терминов	19
Список использованных источников	20

ВВЕДЕНИЕ

Somebody □□□□□□□□□□ once told me [**Article**].

You stupid □□□□

ГЛАВА 1. НАЗВАНИЕ ПЕРВОЙ ГЛАВЫ: ВСЕСТОРОННЕЕ ИЗУЧЕНИЕ ОБЪЕКТА И ПРЕДМЕТА ИССЛЕДОВАНИЯ, АНАЛИЗ РЕЗУЛЬТАТОВ, ПОЛУЧЕННЫХ ДРУГИМИ АВТОРАМИ

Хорошим стилем является наличие введения к главе, которое *начинается непосредственно после названия главы, без оформления в виде отдельного параграфа*. Во введении может быть описана цель написания главы, а также приведена краткая структура главы. Например, в параграфе 1.1 приведены примеры оформления одиночных формул, рисунков и таблицы. Параграф 1.2 посвящён многострочным формулам и сложносоставным рисункам.

Текст данной главы призван привести *краткие* примеры оформления текстово-графических объектов. Более подробные примеры можно посмотреть в следующей главе, а также в рекомендациях студентам [spbpu-student-thesis-template-author-guide].

1.1. Название параграфа

1.1.1. Название первого подпараграфа первого параграфа первой главы для демонстрации переноса слов в содержании

Содержание первого подпараграфа первого параграфа первой главы.

Одиночные формулы оформляют в окружении equation, например, как указано в следующей одиночной нумерованной формуле:

$$\pi \approx 3,141. \tag{1.1}$$

На рис.1.1 изображена гидробашня СПбПУ, а в табл.2.2 приведены данные, на примере которых коротко и наглядно будет изложена суть ВКР.

1.2. Название параграфа

Формулы могут быть размещены в несколько строк. Чтобы выставить номер формулы напротив средней строки, используйте окружение multlined из па-



Рис.1.1. Вид на гидробашню СПбПУ [spbpu-gallery]

кета `mathtools` следующим образом [Ganter1999]:

$$\begin{aligned}
 (A_1, B_1) &\leq (A_2, B_2) \Leftrightarrow \\
 &\Leftrightarrow A_1 \subseteq A_2 \Leftrightarrow \\
 &\Leftrightarrow B_2 \subseteq B_1.
 \end{aligned}
 \tag{1.2}$$

Используя команду `\labelcref` из пакета `cleveref`, допустимо следующим образом оформлять ссылку на несколько формул: (1.1 and 1.2). На рис.1.2 приведены три картинка под общим номером и названием, но с отдельной нумерацией подрисунков посредством пакета `subcaption`.

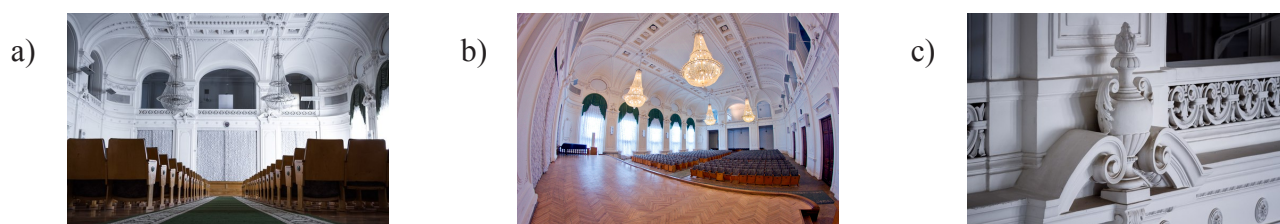


Рис.1.2. Фотографии Белого зала СПбПУ [spbpu-gallery], в том числе: *a* — со стороны зрителей; *b* — со стороны сцены; *c* — барельеф

Далее можно ссылаться на три отдельных рисунка: рис.1.2а, рис.1.2b и рис.1.2с.

Пример ссылок [Article; Book; Booklet; Conference; Inbook; Incollection; Manual; Mastersthesis; Misc; Phdthesis; Proceedings; Techreport; Unpublished; badiou:briefings], а также ссылок с указанием страниц, на котором отображены номера страниц [Naidenova2017] или в виде мультицитаты на несколько источни-

ков [Naidenova2017; Ganter1999]. Часть библиографических записей носит иллюстративный характер и не имеет отношения к реальной литературе.

1.3. Выводы

Текст выводов по главе 1.

Кроме названия параграфа «выводы» можно использовать (единообразно по всем главам) следующие подходы к именованию последних разделов с результатами по главам:

- «выводы по главе N», где N — номер соответствующей главы;
- «резюме»;
- «резюме по главе N», где N — номер соответствующей главы.

Параграф с изложением выводов по главе *является обязательным*.

ГЛАВА 2. НАЗВАНИЕ ВТОРОЙ ГЛАВЫ: РАЗРАБОТКА МЕТОДА, АЛГОРИТМА, МОДЕЛИ ИССЛЕДОВАНИЯ

Глава посвящена более подробным примерам оформления текстово-графических объектов.

В параграфе 2.1 приведены примеры оформления многострочной формулы и одиночного рисунка. Параграф 2.2 раскрывает правила оформления перечислений и псевдокода. В параграфе 2.3 приведены примеры оформления сложно-ставных рисунков, длинных таблиц, а также теоремоподобных окружений.

2.1. Название параграфа

Все формулы, размещенные в отдельных строках, подлежат нумерации, например, как формулы (2.1) и (2.2) из [Ganter1999].

$$A^{\uparrow} = \{m \in M \mid gIm \forall g \in A\}; \quad (2.1)$$

$$B^{\downarrow} = \{g \in G \mid gIm \forall m \in B\}. \quad (2.2)$$

Обратим внимание, что формулы содержат знаки препинания и что они выровнены по левому краю (с помощью знака & окружения align).

На рис.2.1 приведёна фотография Нового научно-исследовательского корпуса СПбПУ.



Рис.2.1. Новый научно-исследовательский корпус СПбПУ [spbpu-gallery]

2.2. Название параграфа

Название параграфа оформляется с помощью команды `\section{...}`, название главы — `\chapter{...}`.

2.2.1. Название подпараграфа

Название подпараграфа оформляется с помощью команды `\subsection{...}`.

Использование подподпараграфов в основной части крайне не рекомендуется. В случае использования, необходимо вынести данный номер в содержание. Название подпараграфа оформляется с помощью команды `\subsubsection{...}`.

Вместо подподпараграфов рекомендовано использовать перечисления.

Перечисления могут быть с нумерационной частью и без неё и использоваться с иерархией и без иерархии. Нумерационная часть при этом формируется следующим способом:

1. в перечислениях *без иерархии* оформляется арабскими цифрами с точкой (или длинным тире).
2. В перечислениях *с иерархией* — в последовательности сначала прописных латинских букв с точкой, затем арабских цифр с точкой и далее — строчных латинских букв со скобкой.

Далее приведён пример перечислений с иерархией.

- A. Первый пункт.
- B. Второй пункт.
- C. Третий пункт.

D. По ГОСТ 2.105–95 [**gost-russian-text-documents**] первый уровень нумерации идёт буквами русского или латинского алфавитов (*для определенности выбираем английский алфавит*), а второй — цифрами.

1. В данном пункте лежит следующий нумерованный список:

- a) первый пункт;
- b) третий уровень нумерации не нормирован ГОСТ 2.105–95 (*для определенности выбираем английский алфавит*);
- c) обращаем внимание на строчность букв в этом нумерованном и следующем маркированном списке:
 - первый пункт маркированного списка.

E. Пятый пункт верхнего уровня перечисления.

Маркированный список (без нумерационной части) используется, если нет необходимости ссылки на определенное положение в списке:

- первый пункт с *маленькой буквы* по правилам русского языка;
- второй пункт с *маленькой буквы* по правилам русского языка.

Оформление псевдокода необходимо осуществлять с помощью пакета `algorithm2e` в окружении `algorithm`. Данное окружение интерпретируется в шаблоне как рисунок. Пример оформления псевдокода алгоритма приведён на рис.2.2.

Обратим внимание, что можно сослаться на строчку 1 псевдокода из рис.2.2.

2.3. Название параграфа

Одиночные формулы также, как и отдельные формулы в составе группы, могут быть размещены в несколько строк. Чтобы выставить номер формулы напротив средней строки, используйте окружение `multlined` из пакета `mathtools` следующим образом [**Ganter1999**]:

$$\begin{aligned}
 (A_1, B_1) &\leq (A_2, B_2) \Leftrightarrow \\
 &\Leftrightarrow A_1 \subseteq A_2 \Leftrightarrow \\
 &\Leftrightarrow B_2 \subseteq B_1.
 \end{aligned}
 \tag{2.3}$$

Используя команду `\labelcref{...}` из пакета `cleveref`, допустимо оформить ссылку на несколько формул, например, (2.1 to 2.3).

Algorithm

Input: the many-valued context $\mathbb{M} \stackrel{\text{def}}{=} (G, M, W, J)$, the class membership $\varepsilon : G \rightarrow K$

Output: positive and negative binary contexts $\overline{\mathbb{K}}_+ \stackrel{\text{def}}{=} (\overline{G}_+, M, I_+)$, $\overline{\mathbb{K}}_- \stackrel{\text{def}}{=} (\overline{G}_-, M, I_-)$ such that i-tests found in $\overline{\mathbb{K}}_+$ are diagnostic tests in \mathbb{M} , and objects from $\overline{\mathbb{K}}_-$ are counter-examples

```

1. for  $\forall g_i, g_j \in G$  do
2.   if  $i < j$  then
3.      $\overline{G} \leftarrow (g_i, g_j);$ 
4.   for  $\forall (g_i, g_j) \in \overline{G}$  do
5.     if  $m(g_i) = m(g_j)$  then
6.        $(g_i, g_j)Im;$ 
7.     if  $\varepsilon(g_i) = \varepsilon(g_j)$  then
8.        $\overline{G}_+ \leftarrow (g_i, g_j);$ 
9.     else  $\overline{G}_- \leftarrow (g_i, g_j);$ 
10.   $I_+ = I \cap (\overline{G}_+ \times M), I_- = I \cap (\overline{G}_- \times M);$ 
11.  for  $\forall \overline{g}_+ \in \overline{G}_+, \forall \overline{g}_- \in \overline{G}_-$  do
12.    if  $\overline{g}_+^\uparrow \subseteq \overline{g}_-^\uparrow$  then
13.       $\overline{G}_+ \leftarrow \overline{G}_+ \setminus \overline{g}_+;$ 

```

Рис.2.2. Псевдокод алгоритма DiagnosticTestsScalingAndInferring [Naidenova2017]

Пример оформления четырёх иллюстраций в одном текстово-графическом объекте приведён на рис.2.3. Это возможно благодаря использованию пакета subcaption.

Далее можно ссылаться на составные части данного рисунка как на самостоятельные объекты: рис.2.3а, рис.2.3b, рис.2.3с, рис.2.3d или на три из четырёх изображений одновременно: рис.2.3а to 2.3с.

Приведём пример табличного представления данных с записью продолжения на следующей странице на табл.2.1.

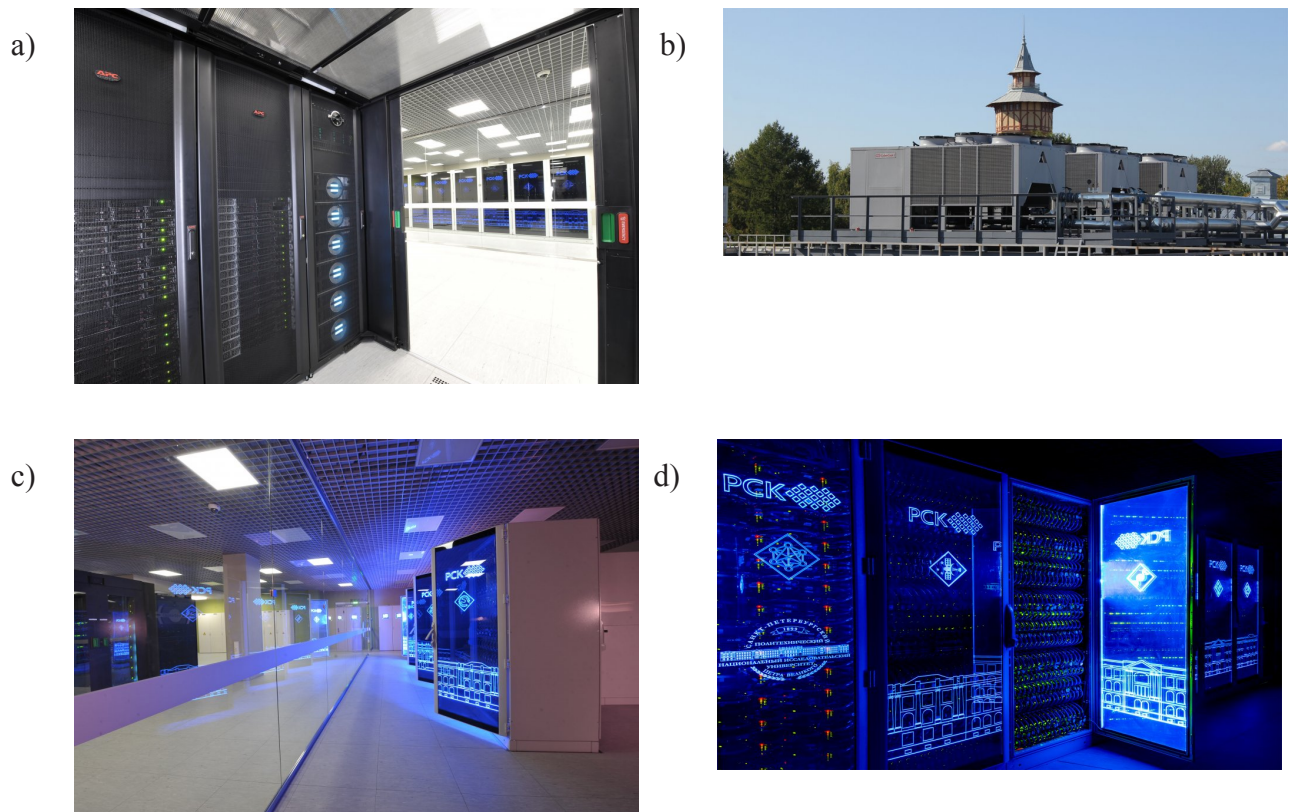


Рис.2.3. Фотографии суперкомпьютерного центра СПбПУ [spbpu-gallery]: *a* — система хранения данных и узлы NUMA-вычислителя; *b* — холодильные машины на крыше научно-исследовательского корпуса; *c* — машинный зал; *d* — элементы вычислительных устройств

Таблица 2.1

Пример задания данных из [Peskov2004] (с повтором для переноса таблицы на новую страницу)

G	m_1	m_2	m_3	m_4	K
1	2	3	4	5	6
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2
g_1	0	1	1	0	1
g_2	1	2	0	1	1

Продолжение табл. 2.1

1	2	3	4	5	6
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2

Таблица 2.2

Пример представления данных для сквозного примера по ВКР [Peskov2004]

G	m_1	m_2	m_3	m_4	K
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2

Таблица 2.3

Пример задания данных в табличном виде из [Peskov2004] (с помощью окружения minipage)

G	m_1	m_2	m_3	m_4	K
g_1	0	1	1	0	1
g_2	1	2	0	1	1
g_3	0	1	0	1	1
g_4	1	2	1	0	2
g_5	1	1	0	1	2
g_6	1	1	1	2	2



Рис.2.4. Новый научно-исследовательский корпус СПбПУ [spbpu-gallery] (с помощью окружения minipage)

Вопросы форматирования текстово-графических объектов (окружений) не регламентированы в известных нам ГОСТах, поэтому предлагаем придерживаться следующих правил:

- **полужирный текст** рекомендуем использовать только для названий стандартных окружений с нумерационной частью, например, для представления *впервые*: **определение 1.1**, **теорема 2.2**, **пример 2.3**, **лемма 4.5**;
- *курсив* рекомендуем использовать только для выделения переменных в формулах, служебной информации об авторах главы (статьи), важных терминов, представляемых по тексту, а также для всего тела окружений, связанных с получением *новых существенных результатов и их доказательством*: теорема, лемма, следствие, утверждение и другие.

По аналогии с нумерацией формул, рисунков и таблиц нумеруются и иные текстово-графические объекты, то есть включаем в нумерацию номер главы, например: теорема 3.1. для первой теоремы третьей главы монографии. Команды L^AT_EX выставляют нумерацию и форматирование автоматически. Полный пере-

чень команд для подготовки текстово-графических и иных объектов находится в подробных методических рекомендациях [**spbpu-bci-template-author-guide**].

Для удобства авторов названия стандартных окружений, рекомендованных к использованию, приведены в табл.2.4, а в табл.2.5 перечислены имена специально разработанных окружений для шаблонов SPbPU.

Таблица 2.4

Стандартные окружения

Название окружения	Назначение
center	центрирование, аналог команды <code>\centering</code> , но с добавлением нежелательного пробела, поэтому лучше избегать применения <code>center</code>
itemize	перечисления, в которых нет необходимости нумеровать пункты (немаркированные списки)
enumerate	перечисления с нумерацией (немаркированные списки)
refsection	создание отдельных библиографических списков для глав
tabular	оформление таблиц
table	автоматическое перемещение по тексту таблиц, оформленных, например, с помощью <code>tabular</code> , для минимизации пустых пространств
longtable	оформление многостраничных таблиц
tikzpicture	создание иллюстраций с помощью пакета <code>tikz</code> [ctan-tikz]
figure	автоматическое перемещение по тексту рисунков, оформленных, например, с помощью <code>tikz</code> или подключенных с помощью команды <code>\includegraphics</code> , для минимизации пустых пространств
subfigure	оформление вложенных рисунков в составе <code>figure</code>
algorithm	оформление псевдокода на основе пакета <code>algorithm2e</code> [ctan-algorithm2e]
minipage	оформление рисунков и таблиц без функций автоматического перемещения по тексту для минимизации пустых пространств
equation	оформление выключенных (не встроенных в текст с помощью <code>\$...\$</code>) одиночных формул на одной строке
multilined	оформление выключенных (не встроенных в текст с помощью <code>\$...\$</code>) одиночных формул в несколько строк
aligned	оформление нескольких формул с выравниванием по символу <code>&</code> .

На базе пакета `tikz` разработано большое количество расширений [**ctan-tikz**], например, `tikzcd`, которые мы рекомендуем использовать для оформления иллюстраций.

В случае, если авторам потребовалось новое окружение, то создать его можно в файле `my_folder/my_settings.tex` согласно правилам, приведённым ниже.

Таблица 2.5

Специальные окружения

Название окружения	Текстово-графический объект
abstr	реферат (abstract)
m-theorem	теорема
m-corollary	следствие
m-proposition	утверждение
m-lemma	лемма
m-axiom	аксиома
m-example	пример
m-definition	определение
m-condition	условие
m-problem	проблема
m-exercise	упражнение
m-question	вопрос
m-hypothesis	гипотеза

1. Для перехода в режим создания окружений следует указать:
 - `\theoremstyle{myplain}` — окружения с доказательствами или аксиомами
 - `\theoremstyle{mydefinition}` — окружения, не связанные с доказательствами или аксиомами.
2. В команде создания окружения следует ввести краткий псевдоним (`m-new-env`) и отображаемое в pdf имя окружения ():
 - `\newtheorem{m-new-env-second}{ }[chapter]`.

Теорема 2.1 (о чем-то конкретном). *Текст теоремы полностью выделен курсивом. Допустимо математические символы не выделять курсивом, если это искажает их значения. Используется абзацный отступ, так как “Абзацы в тексте начинают отступом” в соответствии с ГОСТ 2.105–95. Название теоремы допустимо убрать. Доказательство окончено.*

Доказательство теоремы 2.1, леммы, утверждений, следствий и других подобных окружений (в последнем абзаце) завершаем предложением в котором сказано, что доказательство окончено. Например, доказательство теоремы 2.1 окончено.

Тело доказательства не выделяется курсивом. Тело следующих окружений также не выделяется сплошным курсивом: определение, условие, проблема, пример, упражнение, вопрос, гипотеза и другие.

Определение 2.1 (термин). В тексте определения только *важные термины* выделяются курсивом. Если определение носит лишь вспомогательный характер, то допустимо не использовать окружение `m-definition`, представляя текст определения в обычном абзаце. Ключевые термины при этом обязательно выделяются курсивом.

Вместо теоремо-подобных окружений для вставки небольших текстово-графических объектов иногда используются команды. Типичным примером такого подхода является команда `\footnote{text}`¹, где в аргументе `text` указывают текст *подстрочной ссылки (сноски)*. В них *нельзя добавлять веб-ссылки или цитировать литературу*. Для этих целей используется список литературы. Нумерация сносок сквозная по ВКР без точки на конце выставляется в шаблоне автоматически, однако в каждом приложении к ВКР нумерация, зависящая от номера приложения, выставляется префикс «П», например «П1.1» — первая сноска первого приложения.

2.4. Выводы

Текст заключения ко второй главе. Пример ссылок [**Article**; **Book**; **Booklet**; **Conference**; **Inbook**; **Incollection**; **Manual**; **Mastersthesis**; **Misc**; **Phdthesis**; **Proceedings**; **Techreport**; **Unpublished**; **badiou:briefings**], а также ссылок с указанием страниц, на котором отображены те или иные текстово-графические объекты [**Naidenova2017**] или в виде мультицитаты на несколько источников [**Naidenova2017**; **Ganter1999**]. Часть библиографических записей носит иллюстративный характер и не имеет отношения к реальной литературе.

Короткое имя каждого библиографического источника содержится в специальном файле `my_biblio.bib`, расположенном в папке `my_folder`. Там же находятся исходные данные, которые с помощью программы `Viber` и стилевого файла `Biblatex-GOST` [**ctan-biblatex-gost**] приведены в списке использованных источников согласно ГОСТ 7.0.5-2008. Многообразные реальные примеры исходных библиографических данных можно посмотреть по ссылке [**ctan-biblatex-gost-examples**].

Как правило, ВКР должна состоять из четырех глав. Оставшиеся главы можно создать по образцу первых двух и подключить с помощью команды `\input` к

¹Внимание! Команда вставляется непосредственно после слова, куда вставляется сноска (без пробела). Лишние пробелы также не указываются внутри команды перед и после фигурных скобок.

исходному коду ВКР. Далее в приложении ?? приведены краткие инструкции запуска исходного кода ВКР [**latex-miktex**; **latex-texstudio**].

В приложении ?? приведено подключение некоторых текстово-графических объектов. Они оформляются по приведенным ранее правилам. В качестве номера структурного элемента вместо номера главы используется «П» с номером главы. Текстово-графические объекты из приложений не учитываются в реферате.

ГЛАВА 3. НАЗВАНИЕ ТРЕТЬЕЙ ГЛАВЫ: РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Хорошим стилем является наличие введения к главе. Во введении может быть описана цель написания главы, а также приведена краткая структура главы.

3.1. Название параграфа

3.2. Название параграфа

3.3. Выводы

Текст выводов по главе 3.

ГЛАВА 4. НАЗВАНИЕ ЧЕТВЁРТОЙ ГЛАВЫ. АПРОБАЦИЯ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ, А ИМЕННО: МЕТОДА, АЛГОРИТМА, МОДЕЛИ ИССЛЕДОВАНИЯ

Хорошим стилем является наличие введения к главе. Во введении может быть описана цель написания главы, а также приведена краткая структура главы.

4.1. Название параграфа

4.2. Название параграфа

Пример ссылки на литературу [avtonomova:fyu; Peskov2004-ru; Kotelnikov2004-ru; Kotelnikov2004].

4.3. Выводы

Текст выводов по главе 4.

ЗАКЛЮЧЕНИЕ

Заключение (2 – 5 страниц) обязательно содержит выводы по теме работы, *конкретные предложения и рекомендации* по исследуемым вопросам. Количество общих выводов должно вытекать из количества задач, сформулированных во введении выпускной квалификационной работы.

Предложения и рекомендации должны быть органически увязаны с выводами и направлены на улучшение функционирования исследуемого объекта. При разработке предложений и рекомендаций обращается внимание на их обоснованность, реальность и практическую приемлемость.

Заключение не должно содержать новой информации, положений, выводов и т. д., которые до этого не рассматривались в выпускной квалификационной работе. Рекомендуется писать заключение в виде тезисов.

Последним абзацем в заключении можно выразить благодарность всем людям, которые помогли автору в написании ВКР.

СЛОВАРЬ ТЕРМИНОВ

TeX — язык вёрстки текста и издательская система, разработанные Дональдом Кнутом.

LaTeX — язык вёрстки текста и издательская система, разработанные Лесли Лампортом как надстройка над TeX.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ