*Образец*

*Статья 9-1 Закона РК*

*«Об авторском праве*

*и смежных правах»*

*от 10 июня 1996 года № 6*

**Реферат**

**База данных**

**под названием**

«База данных «HKR Database» русско-казахских рукописных слов и словосочетаний для моделей распознавания текста»

**Фамилия, имя, отчество** *(если оно указано в документе, удостоверяющем личность)* **авторов:**

1. **Нурсеитов Данияр Борисович**
2. **Бостанбеков Кайрат Аратович**
3. **Алимова Анель Нурданбековна**

**Дата создания объекта: 23 сентября 2020 года**

**Область применения**

Для машинного обучения распознавания рукописного текста требуется набор (множество рукописных целевых слов) входных размеченных данных, на которых можно было бы обучить модель. Существует множество датасетов с рукописным текстом, большинство из них содержат текст на английском языке. Набор данных был собран нашей командой собственноручно, так как подобного набора на казахском, русском языках нет в свободном доступе.

**Назначение**

База данных «База данных «HKR Database» русско-казахских рукописных слов и словосочетаний для моделей распознавания текста» предназначена для машинного обучения распознавания рукописного текста, на казахском и руском языках, включающая в себя 3 набора данных.

**Функциональные возможности**

Было собранно 3 набора входных данных.

1. Для первого набора данных был определен набор данных содержащий в себе все названия крупных и мелких городов Казахстана, областей и районов Казахстана, классификация населенных пунктов, сокращенные названия. Были созданы 3 макета форм для заполнения (Приложение 1, Рисунок 1.1–1.3), формы созданы таким образом, чтобы максимально упростить процесс «вырезания» слов из формы. Макеты содержат следующие **области, разделенные горизонтальными и вертикальными линиями:**

Первый макет содержит 56 секторов **печатного текста, содержащие: Фамилия, Имя, Пол, Возраст, Республика, республика, Казахстан, казахстан, РК, Область, область, обл, обл., город, г., Район, район, р-н, Село, село, с., Поселок, поселок, п., Атырауская, Алматинская, Актюбинская, Акмолинская, Жамбылская, Аягоз, Восточно-Казахстанская, Западно-Казахстанская, Южно-Казахстанская, Северо-Казахстанская, Карагандинская, Павлодарская, Кызылординская, Мангистауская, Костанайская, Астана, Абай, Акколь, Аксай, Аксу, Актау, Актобе, Алга, Алматы, Аральск, Аркалык, Арысь, Атбасар, Атырау, Балхаш, Байконур, Булаево, условный код формы, и 56 секторов пустых зон для написания рукописного текста.**

Второй макет содержит 56 секторов **печатного текста, содержащие: Фамилия, Имя, Пол, Возраст, Державинск, Ерейментау, Есик, Есиль, Жанаозен, Жанатас, Жаркент, Жезказган, Жем, Жетысай, Житикара, Зайсан, Зыряновск, Казалинск, Кандыагаш, Капчагай, Караганда, Каражал, Каратау, Каркаралинск, Каскелен, Кентау, Кокшетау, Костанай, Кульсары, Курчатов, Кызылорда, Ленгер, Лисаковск, Макинск, Мамлютка, Павлодар, Петропавловск, Приозерск, Риддер, Рудный, Сарань, Сарканд, Сарыагаш, Сатпаев, Семей, Сергеевка, Серебрянск, Степногорск, Степняк, Тайынша, Талгар, Талдыкорган, Тараз, Текели, Темир, Темиртау, условный код формы, и 56 секторов пустых зон для написания рукописного текста.**

Третий макет содержит 56 секторов **печатного текста, содержащие: Фамилия, Имя, Пол, Возраст, Туркестан, Уральск, Усть-Каменогорск, Ушарал, Уштобе, Форт-Шевченко, Хромтау, Шалкар, Шар, Шардара, Шахтинск, Шемонаиха, Шу, Шымкент, Щучинск, Экибастуз, Эмба, Новодолинский, Айтеке би, Аксу, Аксу, Актас, Актау, Аршалы, Атасу, Балпык, Белколь, Бестобе, Ботакара, Горняцкий, Гульшат, Деркул, Долинка, Жайрем, Жезказган, Заводской, Затобольск, Зачаганск, Индерборский, Карабалык, Качар, Киевка, Конырат, Круглоозерное, Ленинский, Макат, Октябрьский, Осакаровка, Отеген батыра, Сарыколь, Сарыозек, Саяк, условный код формы, и 56 секторов пустых зон для написания рукописного текста.**

2. Для второго набора данных был определен минимальный набор слов, который включает названия городов и областей Казахстана, а также название республик СНГ, и создали макеты форм для заполнения (Приложение 2, Рисунок 2.1). Формы созданы таким образом, чтобы максимально упростить процесс «вырезания» слов из формы

**Форма содержит следующие области, разделенные горизонтальными и вертикальными линиями:**

1. **Фамилия**
2. **Имя**
3. **Пол**
4. **Возраст**
5. **Условный QR-код**
6. **42 сектора печатного текста, содержащие название республик СНГ и названия городов и областей Казахстана. Такие как: Казахстан, Россия, Белоруссия, Украина, Киргизия, Абхазия, Таджикистан, Южная Осетия, Узбекистан, Молдавия, Астана, Акмолинская, Алматы, Алматинская, Актау, Мангистауская, Актобе, Актюбинская, Атырау, Атырауская, Караганда, Карагандинская, Костанай, Костанайская, Кызылорда, Кызылординская, Павлодар, Павлодарская, Петропавловск, Северо-Казахстанская, Семей, Восточно-Казахстанская, Тараз, Жамбылская, Уральск, Западно-Казахстанская, Туркестан, Туркестанская, Кокшетау, Усть-Каменогорск, Шымкент, Южно-Казахстанская**
7. **42 сектора пустых зон для рукописного текста**

3 Для третьего набора данных были отобраны поэмы, включающие различный набор слов русского языка, следующих авторов – классиков:

1. Олжас Сулейменов «Волчата», 1961 г.
2. Олжас Сулейменов «Догони! (Кыз куу)», 1961 г.
3. Михаил Лермонтов «Исповедь», 1831 г.
4. Михаил Лермонтов «Желание», 1831 г.
5. Александр Пушкин «Зимняя дорога», 1826 г.
6. Анна Ахматова «Прошло пять лет», 1950 г.
7. Александр Блок «О доблестях…», 1908 г.
8. Николай Гумилев «Шестое чувство», 1920 г.
9. Иосиф Бродский «Стих», 1980 г.
10. Борис Пастернак «Стога», 1957 г.

Были созданы макеты форм для заполнения (Приложение 3, Рисунок 3.1 – 3.10). Формы созданы таким образом, чтобы максимально упростить процесс «вырезания» слов из формы. Все **макеты разделены горизонтальными и вертикальными линиями.**

**Первая область макета содержит 5 областей: имя, фамилия, название стиха, год стиха, условный QR-код.**

**Вторая область макета содержит 52 сектора печатного текста и 52 сектора пустых зон для рукописного текста для следующих стихов:**

1. Олжас Сулейменов «Волчата», 1961 г. (Приложение 3, Рисунок 3.1) макет содержит следующую информацию:

Шёл человек.

Шёл степью, долго, долго.

Куда? Зачем?

Нам это не узнать.

В густой лощине он увидел волка,

Верней, волчицу,

А, точнее, мать...

Она лежала в зарослях полыни,

Откинув лапы и оскалив пасть.

Из горла перехваченного плыла

Толчками кровь, густая, словно грязь.

Кем? Кем? Волкoм? Охотничьими псами?

Слепым волчатам это не узнать.

Они, толкаясь и ворча, сосали

Большую неподатливую мать.

Голодные волчата позабыли,

Как властно пахнет в зарослях укроп.

Они, прижавшись к маме, жадно пили

Густую холодеющую кровь.

С глотками в них входила жажда мести.

Кому?

Любому.

Лишь бы не простить.

И будут мстить

В отдельности,

Не вместе.

А встретятся –

Друг другу будут мстить.

И человек пошёл своей дорогой.

Куда?.. Зачем?..

Нам это не узнать.

Он был волчатник,

Но волчат не тронул,

Ребят уже не защищала мать...

1. Олжас Сулейменов «Догони! (Кыз куу)», 1961 г. (Приложение 3, Рисунок 3.1) макет содержит следующую информацию:

Догони меня, джигит,

Не жалей коня, джигит.

Если ты влюблён и ловок,

Конь догонит, добежит.

Я люблю тебя, джигит.

Догони же,

Поцелуй,

Голос от стыда дрожит

Среди этих звонких струй.

Меня ветер обгоняет.

На груди моей лежит,

Обнимает, обнимает,

Ой, опять отстал джигит!

Издевается луна,

Я одна.

Опять одна,

Мои руки побелели,

Кровь по крупу скакуна,

Злые люди,

Злые люди,

Вы обидели меня.

Дали смелому джигиту,

И красивому джигиту

Ишака,

А не коня!..

1. Михаил Лермонтов «Исповедь», 1831 г. (Приложение 3, Рисунок 3.3) макет содержит следующую информацию:

Я верю, обещаю верить,

Хоть сам того не испытал,

Что мог монах не лицемерить

И жить, как клятвой обещал;

Что поцелуи и улыбки

Людей коварны не всегда,

Что ближних малые ошибки

Они прощают иногда,

Что время лечит от страданья,

Что мир для счастья сотворен,

Что добродетель не названье

И жизнь поболее, чем сон!..

Но вере теплой опыт хладный

Противуречит каждый миг,

И ум, как прежде безотрадный,

Желанной цели не достиг;

И сердце, полно сожалений,

Хранит в себе глубокий след

Умерших — но святых видений,

И тени чувств, каких уж нет;

Его ничто не испугает,

И то, что было б яд другим,

Его живит, его питает

Огнем язвительным своим.

1. Михаил Лермонтов «Желание», 1831 г. (Приложение 3, Рисунок 3.4) макет содержит следующую информацию:

Зачем я не птица, не ворон степной,

Пролетевший сейчас надо мной?

Зачем не могу в небесах я парить

И одну лишь свободу любить?

На запад, на запад помчался бы я,

Где цветут моих предков поля,

Где в замке пустом, на туманных горах,

Их забвенный покоится прах.

На древней стене их наследственный щит

И заржавленный меч их висит.

Я стал бы летать над мечом и щитом,

И смахнул бы я пыль с них крылом;

И арфы шотландской струну бы задел,

И по сводам бы звук полетел;

Внимаем одним, и одним пробуждён,

Как раздался, так смолкнул бы он.

Но тщетны мечты, бесполезны мольбы

Против строгих законов судьбы.

Меж мной и холмами отчизны моей

Расстилаются волны морей.

Последний потомок отважных бойцов

Увядает средь чуждых снегов;

Я здесь был рожден, но нездешний душой...

О! зачем я не ворон степной?..

1. Александр Пушкин «Зимняя дорога», 1826 г. (Приложение 3, Рисунок 3.5) макет содержит следующую информацию:

Сквозь волнистые туманы

Пробирается луна,

На печальные поляны

Льет печально свет она.

По дороге зимней, скучной

Тройка борзая бежит,

Колокольчик однозвучный

Утомительно гремит.

Что-то слышится родное

В долгих песнях ямщика:

То разгулье удалое,

То сердечная тоска...

Ни огня, ни черной хаты,

Глушь и снег... Навстречу мне

Только версты полосаты

Попадаются одне...

Скучно, грустно... Завтра, Нина,

Завтра к милой возвратясь,

Я забудусь у камина,

Загляжусь не наглядясь.

Звучно стрелка часовая

Мерный круг свой совершит,

И, докучных удаляя,

Полночь нас не разлучит.

Грустно, Нина: путь мой скучен,

Дремля смолкнул мой ямщик,

Колокольчик однозвучен,

Отуманен лунный лик.

1. Анна Ахматова «Прошло пять лет», 1950 г. (Приложение 3, Рисунок 3.6) макет содержит следующую информацию:

Прошло пять лет, – и залечила раны,

Жестокой нанесенные войной,

Страна моя, и русские поляны

Опять полны студеной тишиной.

И маяки сквозь мрак приморской ночи,

Путь указуя моряку, горят.

На их огонь, как в дружеские очи,

Далеко с моря моряки глядят.

Где танк гремел — там ныне мирный трактор,

Где выл пожар — благоухает сад,

И по изрытому когда-то тракту

Автомобили легкие летят.

Где елей искалеченные руки

Взывали к мщенью — зеленеет ель,

И там, где сердце ныло от разлуки, –

Там мать поет, качая колыбель.

Ты стала вновь могучей и свободной,

Страна моя! Но живы навсегда

В сокровищнице памяти народной

Войной испепеленные года.

Для мирной жизни юных поколений,

От Каспия и до полярных льдов,

Как памятники выжженных селений,

Встают громады новых городов.

1. Александр Блок «О доблестях…», 1908 г. (Приложение 3, Рисунок 3.7) макет содержит следующую информацию:

О доблестях, о подвигах, о славе

Я забывал на горестной земле,

Когда твое лицо в простой оправе

Перед мной сияло на столе.

Но час настал, и ты ушла из дому.

Я бросил в ночь заветное кольцо.

Ты отдала свою судьбу другому,

И я забыл прекрасное лицо.

Летели дни, крутясь проклятым роем...

Вино и страсть терзали жизнь мою...

И вспомнил я тебя пред аналоем,

И звал тебя, как молодость свою...

Я звал тебя, но ты не оглянулась,

Я слезы лил, но ты не снизошла.

Ты в синий плащ печально завернулась,

В сырую ночь ты из дому ушла.

Не знаю, где приют твоей гордыне

Ты, милая, ты, нежная, нашла...

Я крепко сплю, мне снится плащ твой синий,

В котором ты в сырую ночь ушла...

Уж не мечтать о нежности, о славе,

Все миновалось, молодость прошла!

Твое лицо в его простой оправе

Своей рукой убрал я со стола.

1. Николай Гумилев «Шестое чувство», 1920 г. (Приложение 3, Рисунок 3.8) макет содержит следующую информацию:

Прекрасно в нас влюбленное вино

И добрый хлеб, что в печь для нас садится,

И женщина, которою дано,

Сперва измучившись, нам насладиться.

Но что нам делать с розовой зарей

Над холодеющими небесами,

Где тишина и неземной покой,

Что делать нам с бессмертными стихами?

Ни съесть, ни выпить, ни поцеловать.

Мгновение бежит неудержимо,

И мы ломаем руки, но опять

Осуждены идти всё мимо, мимо.

Как мальчик, игры позабыв свои,

Следит порой за девичьим купаньем

И, ничего не зная о любви,

Все ж мучится таинственным желаньем;

Как некогда в разросшихся хвощах

Ревела от сознания бессилья

Тварь скользкая, почуя на плечах

Еще не появившиеся крылья;

Так век за веком — скоро ли, Господь? –

Под скальпелем природы и искусства

Кричит наш дух, изнемогает плоть,

Рождая орган для шестого чувства.

1. Иосиф Бродский «Стих», 1980 г. (Приложение 3, Рисунок 3.9) макет содержит следующую информацию:

Я входил вместо дикого зверя в клетку,

выжигал свой срок и кликуху гвоздем в бараке,

жил у моря, играл в рулетку,

обедал черт знает с кем во фраке.

С высоты ледника я озирал полмира,

трижды тонул, дважды бывал распорот.

Бросил страну, что меня вскормила.

Из забывших меня можно составить город.

Я слонялся в степях, помнящих вопли гунна,

надевал на себя что сызнова входит в моду,

сеял рожь, покрывал черной толью гумна

и не пил только сухую воду.

Я впустил в свои сны вороненый зрачок конвоя,

жрал хлеб изгнанья, не оставляя корок.

Позволял своим связкам все звуки, помимо воя;

перешел на шепот. Теперь мне сорок.

Что сказать мне о жизни? Что оказалась длинной.

Только с горем я чувствую солидарность.

Но пока мне рот не забили глиной,

из него раздаваться будет лишь благодарность.

1. Борис Пастернак «Стога», 1957 г. (Приложение 3, Рисунок 3.10) макет содержит следующую информацию:

Снуют пунцовые стрекозы,

Летят шмели во все концы,

Колхозники смеются с возу,

Проходят с косами косцы.

Пока хорошая погода,

Гребут и ворошат корма

И складывают до захода

В стога, величиной с дома.

Стог принимает на закате

Вид постоялого двора,

Где [ночь](http://www.ruthenia.ru/60s/hypertext/noch.htm) ложится на полати

В накошенные клевера.

К утру, когда потемки реже,

Стог высится, как сеновал,

В котором месяц мимоезжий,

Зарывшись, переночевал.

Чем свет телега за телегой

Лугами катятся впотьмах.

Наставший день встает с ночлега

С трухой и сеном в волосах.

А в полдень вновь синеют выси,

Опять стога, как облака,

Опять, как водка на анисе,

Земля душистая крепка.

Форма разработана таким образом, чтобы можно было легко идентифицировать и сегментировать по ячейкам. Для идентификации формы есть маркер в правом верхнем углу каждой формы. Чтобы упростить процесс сегментации, вся форма разделена горизонтальными и вертикальными линиями, что позволяет легко восстановить структуру документа и соответственно пространственное положение слова. Слова индексируются (аннотируются) в соответствии с их положением в таблице. Чтобы вырезать клетки из формы выполняются следующие действия (предварительная обработка):

* сортировка форм по у**словному QR-коду** (маркеру);
* фильтрация формы для улучшения границ таблицы;
* определение контуров таблицы;
* определение и компенсация угла поворота;
* вырезание столбцов;
* вырезание строк;
* разделение строк на слова;
* именование и сохранение слов;

После того, как области изображения, соответствующие ячейкам слова, были сегментированы, данные области изображения могут содержать некоторые краевые артефакты. Например, линий, вырезанные из ячейки или части слова из соседней ячейки можно отнести к артефактам (Рисунок 1). Эти артефакты находятся путем построения вертикальных и горизонтальных гистограмм, затем отрезаются лишние края и очищаются горизонтальные линий.



D:\Projects\KazPost\KazPost Database\Dataets_old_segmentation_versions\cropped_poems\0_1_834.jpg

Рисунок 1 – Пример слов с артефактами

Однако не всегда удается устранить все артефакты. Ниже перечислены некоторые аспекты, которые затрудняют дальнейшую обработку сегментированного слова:

* Расположение букв и их размер существенно различаются от слова к слову.
* Слова могут быть написаны разными цветами (синий, черный, красный, зеленый).
* Слова могут быть зачеркнуты и исправлены.
* В одной ячейке фраза может быть написано в двух строках.

В связи с этим мы разработали систему рекомендаций что позволяет упростить процесс выбора участков слова из формы.

* Предлагаем заполнять форму синей ручкой. Это позволит системе отличать слово от границ таблицы на цветовом уровне. Например, преобразовав изображение из RGB в HSV, мы получаем цветовое представление объектов, инвариантное относительно освещения. В этом цветовом пространстве синий остается синим, независимо от яркости и интенсивности изображения.
* Иногда удаление частей слов из соседних ячеек невозможно без искажения целевого контента данной ячейки; поэтому при заполнении формы желательно, чтобы предмет не выходил за пределы клетки.

Для сегментации слов из таблицы были выполнены следующие шаги:

* Найдена интересующая область в форме – это две колонки, которые заполняются писателями, и разделена в два изображения.
* Сегментированы строки в зависимости от горизонтального белого просвета между строками с помощью построенной гистограммы (Рисунок 2).

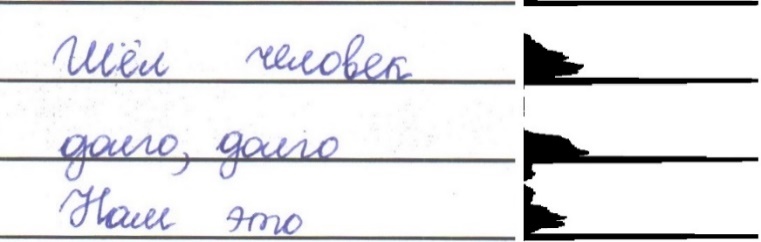


Рисунок 2 – Горизонтальная гистограмма

* В вырезанной строке очистили черную горизонтальную линию от ячейки (Рисунок 3).



Рисунок 3 – Пример с черной горизонтальной линией

* Затем, мы обрезали лишние боковые края слова в зависимости от вертикальной гистограммы (Рисунок 4).

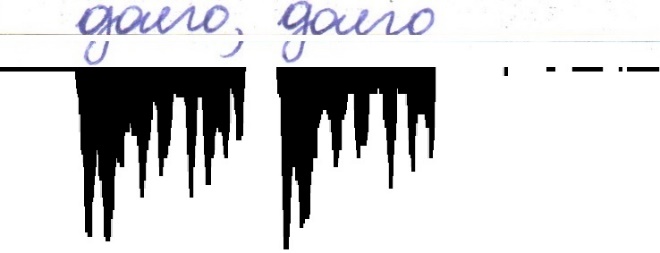


Рисунок 4 – Вертикальная гистограмма

* В результате получили окончательное изображение (Рисунок 5), и сохранили в нашем датасете вместе с аннотацией.

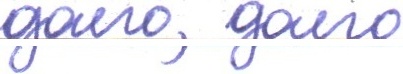


Рисунок 5 – Окончательное изображение

**Основные технические характеристики:**

Название базы данных: **HKR\_Dataset\_Forms\_Public**

* Объем базы данных: 1,91 ГБ.
* Тип файлов: изображения в формате \*.jpg и \*.png
* Количество файлов: 14 папок и 1577 файлов.
* Min размер файла в базе данных: 783 Кбайт
* Max размер файла в базе данных: 4.41 Mбайт

Для соблюдения конфиденциальности, в ф**ормах где содержатся данные: фамилия, имя, пол, возраст и условный код, данные были удалены. В таблице 1 имеются данные по количеству изображении заполненных форм. Пример заполненной формы в Приложении 4, Рисунок 4.1.**

Таблица 1 – Данные по количеству заполненных форм

|  |  |  |
| --- | --- | --- |
| № | Наименование формы | Количество заполненных форм |
|  | 1 Форма, содержащая в себе названия мелких и крупных городов Казахстана | 71 форм |
|  | 2 Форма, содержащая в себе названия мелких и крупных городов Казахстана | 70 форм |
|  | 3 Форма, содержащая в себе названия мелких и крупных городов Казахстана | 70 форм |
|  | Форма, содержащая в себе названия городов и областей Казахстана, а также название республик СНГ | 477 форм |
|  | Форма, содержащая в себе стих Олжаса Сулейменова «Волчата», 1961 г. (Приложение 3, Рисунок 3.1) | 95 форм |
|  | Форма, содержащая в себе стих Олжаса Сулейменова «Догони! (Кыз куу)», 1961 г. (Приложение 3, Рисунок 3.2) | 92 формы |
|  | Форма, содержащая в себе стих Михаила Лермонтова «Исповедь», 1831 г. (Приложение 3, Рисунок 3.3) | 91 форм |
|  | Форма, содержащая в себе стих Михаила Лермонтова «Желание», 1831 г. (Приложение 3, Рисунок 3.4) | 91 форм |
|  | Форма, содержащая в себе стих Александра Пушкина «Зимняя дорога», 1826 г. (Приложение 3, Рисунок 3.5) | 91 форм |
|  | Форма, содержащая в себе стих Анны Ахматовой «Прошло пять лет», 1950 г. (Приложение 3, Рисунок 3.6) | 86 форм |
|  | Форма, содержащая в себе стих Александра Блока «О доблестях…», 1908 г. (Приложение 3, Рисунок 3.7) | 87 форм |
|  | Форма, содержащая в себе стих Николая Гумилева «Шестое чувство», 1920 г. (Приложение 3, Рисунок 3.8) | 84 формы |
|  | Форма, содержащая в себе стих Иосифа Бродского «Стих», 1980 г. (Приложение 3, Рисунок 3.9) | 87 форм |
|  | Форма, содержащая в себе стих Бориса Пастернака «Стога», 1957 г. (Приложение 3, Рисунок 3.10) | 85 форм |

* Название базы данных с аннотациями: **HKR\_Dataset\_Words\_Public**
* Объем базы данных: 721 MБ.
* Тип файлов: 1 папка с изображениями в формате \*.jpg и 1 папка с аннотациями в формате \*.json
* Количество файлов: 2 папки: в одной папке находятся 64943 файлов с изображениями, во второй папке аннотации к этим изображениям в количестве 64943 файлов.
* Min размер файла в базе данных: в папке с изображениями 1,59 КБ, в папке с аннотациями 140 байт
* Max размер файла в базе данных: в папке с изображениями 38,3 КБ, в папке с аннотациями 222 байта

**В таблице 2 имеются данные по количеству обработанных изображении с аннотациями.**

Таблица 2 – Данные по количеству заполненных форм

| № | Слово или словосочетание | Количество изображений |
| --- | --- | --- |
|  | Шёл человек. | 88 |
|  | долго, долго. | 81 |
|  | Нам это | 80 |
|  | В густой лощине | 78 |
|  | Верней, волчицу, | 74 |
|  | Она лежала в | 78 |
|  | Откинув лапы | 76 |
|  | Из горла | 69 |
|  | Толчками кровь, | 69 |
|  | Кем? Кем? Волкoм? | 79 |
|  | Слепым волчатам | 86 |
|  | Они, толкаясь | 83 |
|  | Большую | 75 |
|  | Голодные волчата | 73 |
|  | Как властно пахнет | 73 |
|  | Они, прижавшись | 80 |
|  | Густую холодеющую | 70 |
|  | С глотками в них | 75 |
|  | Кому? Любому. | 74 |
|  | И будут мстить | 72 |
|  | Не вместе. | 75 |
|  | Друг другу | 69 |
|  | И человек пошёл | 69 |
|  | Куда?.. Зачем?.. | 74 |
|  | Он был волчатник, | 71 |
|  | Ребят уже не | 64 |
|  | Шёл степью, | 85 |
|  | Куда? Зачем? | 79 |
|  | не узнать. | 70 |
|  | он увидел волка, | 64 |
|  | А, точнее, мать... | 72 |
|  | зарослях полыни, | 80 |
|  | и оскалив пасть. | 81 |
|  | перехваченного плыла | 83 |
|  | густая, словно грязь. | 79 |
|  | Охотничьими псами? | 80 |
|  | это не узнать. | 76 |
|  | и ворча, сосали | 75 |
|  | неподатливую мать. | 79 |
|  | позабыли, | 77 |
|  | в зарослях укроп. | 76 |
|  | к маме, жадно пили | 75 |
|  | кровь. | 76 |
|  | входила жажда мести. | 69 |
|  | Лишь бы не простить. | 71 |
|  | В отдельности, | 77 |
|  | А встретятся – | 75 |
|  | будут мстить. | 71 |
|  | своей дорогой. | 69 |
|  | Нам это не узнать. | 69 |
|  | Но волчат не тронул, | 69 |
|  | защищала мать... | 73 |
|  | Догони меня, | 86 |
|  | Не жалей коня, | 82 |
|  | Если | 78 |
|  | влюблён | 84 |
|  | Конь догонит, | 73 |
|  | Я люблю тебя, | 75 |
|  | Догони же, | 85 |
|  | Голос | 84 |
|  | стыда | 77 |
|  | Среди этих | 69 |
|  | Меня ветер | 70 |
|  | На груди | 73 |
|  | Обнимает, | 74 |
|  | Ой, | 80 |
|  | отстал | 89 |
|  | Издевается | 80 |
|  | Я одна. | 76 |
|  | Мои руки | 70 |
|  | Кровь | 74 |
|  | крупу | 76 |
|  | Злые | 76 |
|  | Злые | 79 |
|  | Вы обидели | 73 |
|  | Дали смелому | 67 |
|  | И красивому | 65 |
|  | Ишака, | 60 |
|  | джигит, | 85 |
|  | джигит. | 83 |
|  | ты | 76 |
|  | и ловок, | 87 |
|  | добежит. | 86 |
|  | джигит. | 73 |
|  | Поцелуй, | 76 |
|  | от | 69 |
|  | дрожит | 68 |
|  | звонких струй. | 67 |
|  | обгоняет. | 75 |
|  | моей лежит, | 87 |
|  | обнимает, | 87 |
|  | опять | 92 |
|  | джигит! | 85 |
|  | луна, | 80 |
|  | Опять одна, | 73 |
|  | побелели, | 73 |
|  | по | 45 |
|  | скакуна, | 83 |
|  | люди, | 79 |
|  | люди, | 85 |
|  | меня. | 79 |
|  | джигиту, | 89 |
|  | джигиту | 62 |
|  | А не коня!.. | 57 |
|  | Я верю, | 76 |
|  | Хоть сам того | 75 |
|  | Что мог монах | 87 |
|  | И жить, как | 84 |
|  | Что поцелуи | 67 |
|  | Людей коварны | 67 |
|  | Что ближних | 74 |
|  | Они прощают | 74 |
|  | Что время лечит | 72 |
|  | Что мир для | 64 |
|  | Что добродетель | 64 |
|  | И жизнь поболее, | 66 |
|  | Но вере | 74 |
|  | опыт | 86 |
|  | Противуречит | 67 |
|  | И ум, как прежде | 63 |
|  | Желанной цели | 68 |
|  | И сердце, | 71 |
|  | Хранит в себе | 70 |
|  | Умерших, но | 72 |
|  | И тени чувств, | 67 |
|  | Его ничто | 71 |
|  | И то, | 82 |
|  | было б | 83 |
|  | Его живит, | 84 |
|  | Огнем | 81 |
|  | обещаю верить, | 83 |
|  | не испытал, | 85 |
|  | не лицемерить | 75 |
|  | клятвой обещал; | 73 |
|  | и улыбки | 70 |
|  | не всегда, | 70 |
|  | малые ошибки | 75 |
|  | иногда, | 85 |
|  | от страданья, | 73 |
|  | счастья сотворен, | 71 |
|  | не названье | 75 |
|  | чем сон!.. | 77 |
|  | теплой | 87 |
|  | хладный | 79 |
|  | каждый миг, | 71 |
|  | безотрадный, | 71 |
|  | не достиг; | 76 |
|  | полно сожалений, | 77 |
|  | глубокий след | 54 |
|  | святых видений – | 53 |
|  | каких уж нет; | 81 |
|  | не испугает, | 76 |
|  | что | 71 |
|  | яд другим, | 75 |
|  | его питает | 70 |
|  | язвительным своим. | 71 |
|  | Зачем я не птица, | 79 |
|  | Пролетевший | 76 |
|  | Зачем не могу | 74 |
|  | И одну лишь | 68 |
|  | На запад, на запад | 60 |
|  | Где цветут моих | 60 |
|  | Где в замке пустом, | 67 |
|  | Их забвенный | 67 |
|  | На древней стене | 74 |
|  | И заржавленный | 76 |
|  | Я стал | 78 |
|  | над мечом | 66 |
|  | И смахнул бы я | 65 |
|  | И арфы шотландской | 65 |
|  | И по сводам бы | 65 |
|  | Внимаем одним, | 71 |
|  | Как раздался, | 66 |
|  | Но тщетны мечты, | 71 |
|  | Против строгих | 76 |
|  | Меж мной и холмами | 77 |
|  | Расстилаются | 83 |
|  | Последний потомок | 69 |
|  | Увядает средь | 68 |
|  | Я здесь | 64 |
|  | но нездешний | 64 |
|  | О! зачем я | 60 |
|  | не ворон степной, | 83 |
|  | сейчас надо мной? | 73 |
|  | в небесах я парить | 72 |
|  | свободу любить? | 70 |
|  | помчался бы я, | 76 |
|  | предков поля, | 78 |
|  | на туманных горах, | 78 |
|  | покоится прах. | 81 |
|  | их наследственный щит | 79 |
|  | меч их висит. | 81 |
|  | бы летать | 84 |
|  | и щитом, | 84 |
|  | пыль с них крылом; | 80 |
|  | струну бы задел, | 76 |
|  | звук полетел; | 73 |
|  | и одним пробуждён, | 72 |
|  | так смолкнул бы он. | 71 |
|  | бесполезны мольбы | 71 |
|  | законов судьбы. | 77 |
|  | отчизны моей | 75 |
|  | волны морей. | 72 |
|  | отважных бойцов | 70 |
|  | чуждых снегов; | 74 |
|  | был рожден, | 76 |
|  | душой... | 53 |
|  | не ворон степной?.. | 42 |
|  | Сквозь волнистые | 71 |
|  | Пробирается | 66 |
|  | На печальные | 71 |
|  | Льет печально | 82 |
|  | По дороге зимней, | 66 |
|  | Тройка борзая | 66 |
|  | Колокольчик | 74 |
|  | Утомительно | 85 |
|  | Что-то слышится | 85 |
|  | В долгих | 67 |
|  | То разгулье удалое, | 59 |
|  | Ни огня, | 65 |
|  | Глушь и снег... | 65 |
|  | Только версты полосаты | 56 |
|  | Скучно, грустно... | 62 |
|  | Завтра к милой | 62 |
|  | Я забудусь | 68 |
|  | Загляжусь | 60 |
|  | Звучно стрелка | 61 |
|  | Мерный круг | 69 |
|  | И, докучных | 68 |
|  | Полночь нас | 64 |
|  | Грустно, Нина: | 65 |
|  | Дремля смолкнул | 65 |
|  | Колокольчик | 71 |
|  | Отуманен | 59 |
|  | туманы | 73 |
|  | луна, | 72 |
|  | поляны | 76 |
|  | свет она. | 80 |
|  | скучной | 70 |
|  | бежит, | 67 |
|  | однозвучный | 75 |
|  | гремит. | 78 |
|  | родное | 78 |
|  | песнях ямщика: | 74 |
|  | То сердечная тоска... | 76 |
|  | ни черной хаты, | 74 |
|  | Навстречу мне | 65 |
|  | Попадаются одне... | 63 |
|  | Завтра, Нина, | 66 |
|  | возвратясь, | 68 |
|  | у камина, | 73 |
|  | не наглядясь. | 65 |
|  | часовая | 65 |
|  | свой совершит, | 75 |
|  | удаляя, | 67 |
|  | не разлучит. | 61 |
|  | путь мой скучен, | 64 |
|  | мой ямщик, | 66 |
|  | однозвучен, | 61 |
|  | лунный лик. | 48 |
|  | Прошло пять лет, – | 69 |
|  | Жестокой | 71 |
|  | Страна моя, | 70 |
|  | Опять полны | 70 |
|  | И маяки сквозь | 72 |
|  | Путь указуя | 67 |
|  | На их огонь, | 69 |
|  | Далеко с моря | 70 |
|  | Где танк | 66 |
|  | там ныне | 73 |
|  | Где выл пожар — | 69 |
|  | И по изрытому | 63 |
|  | Автомобили | 65 |
|  | Где елей | 60 |
|  | Взывали к мщенью — | 55 |
|  | И там, где сердце | 60 |
|  | Там мать поет, | 65 |
|  | Ты стала вновь | 74 |
|  | Страна | 66 |
|  | Но живы | 68 |
|  | В сокровищнице | 69 |
|  | Войной | 71 |
|  | Для мирной жизни | 63 |
|  | От Каспия и | 63 |
|  | Как памятники | 70 |
|  | Встают громады | 57 |
|  | и залечила раны, | 71 |
|  | нанесенные войной, | 70 |
|  | и русские поляны | 69 |
|  | студеной тишиной. | 67 |
|  | мрак приморской ночи, | 69 |
|  | моряку, горят. | 68 |
|  | как в дружеские очи, | 67 |
|  | моряки глядят. | 74 |
|  | гремел — | 74 |
|  | мирный трактор, | 71 |
|  | благоухает сад, | 69 |
|  | когда-то тракту | 70 |
|  | легкие летят. | 76 |
|  | искалеченные руки | 76 |
|  | зеленеет ель, | 72 |
|  | ныло от разлуки, – | 71 |
|  | качая колыбель. | 69 |
|  | могучей и свободной, | 69 |
|  | моя! | 69 |
|  | навсегда | 66 |
|  | памяти народной | 63 |
|  | испепеленные года. | 69 |
|  | юных поколений, | 68 |
|  | до полярных льдов, | 61 |
|  | выжженных селений, | 56 |
|  | новых городов. | 54 |
|  | О доблестях, | 67 |
|  | Я забывал на | 61 |
|  | Когда твое лицо | 64 |
|  | Перед мной | 67 |
|  | Но час настал, | 73 |
|  | Я бросил в ночь | 71 |
|  | Ты отдала свою | 65 |
|  | И я забыл | 68 |
|  | Летели дни, крутясь | 55 |
|  | Вино и страсть | 59 |
|  | И вспомнил я тебя | 70 |
|  | И звал | 79 |
|  | как молодость | 72 |
|  | Я звал тебя, | 68 |
|  | Я слезы лил, | 62 |
|  | Ты в синий плащ | 65 |
|  | В сырую ночь | 68 |
|  | Не знаю, где приют | 61 |
|  | Ты, милая, ты, | 66 |
|  | Я крепко | 76 |
|  | мне снится плащ | 75 |
|  | В котором ты | 72 |
|  | Уж не мечтать | 73 |
|  | Все миновалось, | 74 |
|  | Твое лицо в его | 67 |
|  | Своей рукой убрал | 51 |
|  | о подвигах, о славе | 72 |
|  | горестной земле, | 70 |
|  | в простой оправе | 69 |
|  | сияло на столе. | 76 |
|  | и ты ушла из дому. | 66 |
|  | заветное кольцо. | 66 |
|  | судьбу другому, | 70 |
|  | прекрасное лицо. | 73 |
|  | проклятым роем... | 76 |
|  | терзали жизнь мою... | 72 |
|  | пред аналоем, | 69 |
|  | тебя, | 70 |
|  | свою... | 79 |
|  | но ты не оглянулась, | 78 |
|  | но ты не снизошла. | 75 |
|  | печально завернулась, | 75 |
|  | ты из дому ушла. | 65 |
|  | твоей гордыне | 63 |
|  | нежная, нашла... | 72 |
|  | сплю, | 77 |
|  | твой синий, | 78 |
|  | в сырую ночь ушла... | 67 |
|  | о нежности, о славе, | 66 |
|  | молодость прошла! | 67 |
|  | простой оправе | 64 |
|  | я со стола. | 71 |
|  | Прекрасно в нас | 69 |
|  | И добрый | 62 |
|  | что в печь | 67 |
|  | И женщина, | 69 |
|  | Сперва измучившись, | 63 |
|  | Но что нам делать | 67 |
|  | Над холодеющими | 64 |
|  | Где тишина | 60 |
|  | Что делать нам с | 66 |
|  | Ни съесть, ни выпить, | 66 |
|  | Мгновение бежит | 67 |
|  | И мы ломаем руки, | 68 |
|  | Осуждены идти | 60 |
|  | Как мальчик, | 65 |
|  | Следит порой за | 70 |
|  | И, ничего | 71 |
|  | Все ж мучится | 66 |
|  | Как некогда в | 61 |
|  | Ревела от | 67 |
|  | Тварь скользкая, | 62 |
|  | Еще не появившиеся | 61 |
|  | Так век | 74 |
|  | скоро ли, | 71 |
|  | Под скальпелем | 60 |
|  | Кричит наш дух, | 54 |
|  | Рождая орган для | 41 |
|  | влюбленное вино | 71 |
|  | хлеб, | 71 |
|  | для нас садится, | 74 |
|  | которою дано, | 71 |
|  | нам насладиться. | 64 |
|  | с розовой зарей | 66 |
|  | небесами, | 69 |
|  | и неземной покой, | 67 |
|  | бессмертными стихами? | 70 |
|  | ни поцеловать. | 74 |
|  | неудержимо, | 77 |
|  | но опять | 76 |
|  | всё мимо, мимо. | 77 |
|  | игры позабыв свои, | 68 |
|  | девичьим купаньем | 62 |
|  | не зная о любви, | 69 |
|  | таинственным желаньем; | 73 |
|  | разросшихся хвощах | 67 |
|  | сознания бессилья | 64 |
|  | почуя на плечах | 74 |
|  | крылья; | 70 |
|  | за веком — | 67 |
|  | Господь? – | 57 |
|  | природы и искусства | 57 |
|  | изнемогает плоть, | 55 |
|  | шестого чувства. | 52 |
|  | Я входил вместо | 62 |
|  | выжигал свой | 62 |
|  | и кликуху гвоздем | 69 |
|  | жил у моря, | 63 |
|  | обедал черт знает | 60 |
|  | С высоты ледника | 63 |
|  | трижды тонул, | 66 |
|  | Бросил страну, | 59 |
|  | Из забывших меня | 62 |
|  | Я слонялся в степях, | 62 |
|  | надевал на себя что | 63 |
|  | сеял рожь, покрывал | 66 |
|  | и не пил только | 73 |
|  | Я впустил | 67 |
|  | вороненый зрачок | 66 |
|  | жрал хлеб | 69 |
|  | не оставляя | 74 |
|  | Позволял своим | 67 |
|  | все звуки, | 65 |
|  | перешел на шепот. | 70 |
|  | Что сказать мне | 70 |
|  | Что оказалась | 60 |
|  | Только с горем я | 58 |
|  | Но пока мне рот | 65 |
|  | из него | 62 |
|  | будет лишь | 52 |
|  | дикого зверя в клетку, | 72 |
|  | срок | 68 |
|  | в бараке, | 65 |
|  | играл в рулетку, | 60 |
|  | с кем во фраке. | 64 |
|  | я озирал полмира, | 67 |
|  | дважды бывал распорот. | 60 |
|  | что меня вскормила. | 62 |
|  | можно составить город. | 63 |
|  | помнящих вопли гунна, | 57 |
|  | сызнова входит в моду, | 55 |
|  | черной толью гумна | 60 |
|  | сухую воду. | 60 |
|  | в свои сны | 63 |
|  | конвоя, | 77 |
|  | изгнанья, | 76 |
|  | корок. | 68 |
|  | связкам | 69 |
|  | помимо воя; | 67 |
|  | Теперь мне сорок. | 68 |
|  | о жизни? | 68 |
|  | длинной. | 58 |
|  | чувствую солидарность. | 55 |
|  | не забили глиной, | 61 |
|  | раздаваться | 61 |
|  | благодарность. | 44 |
|  | Снуют пунцовые | 73 |
|  | Летят шмели | 72 |
|  | Колхозники | 70 |
|  | Проходят с | 66 |
|  | Пока хорошая | 65 |
|  | Гребут и | 63 |
|  | И складывают | 65 |
|  | В стога, | 68 |
|  | Стог принимает | 73 |
|  | Вид постоялого | 67 |
|  | Где [ночь](http://www.ruthenia.ru/60s/hypertext/noch.htm) | 65 |
|  | В накошенные | 71 |
|  | К утру, когда | 62 |
|  | Стог высится, | 61 |
|  | В котором | 69 |
|  | Зарывшись, | 63 |
|  | Чем свет телега | 67 |
|  | Лугами катятся | 68 |
|  | Наставший | 69 |
|  | встает с | 80 |
|  | С трухой и | 60 |
|  | А в | 59 |
|  | вновь синеют | 78 |
|  | Опять стога, | 78 |
|  | Опять, как | 73 |
|  | Земля душиста | 54 |
|  | стрекозы, | 65 |
|  | во все концы, | 64 |
|  | смеются с возу, | 74 |
|  | косами косцы. | 77 |
|  | погода, | 63 |
|  | ворошат корма | 63 |
|  | до захода | 63 |
|  | величиной с дома. | 63 |
|  | на закате | 70 |
|  | двора, | 68 |
|  | ложится на полати | 75 |
|  | клевера. | 77 |
|  | потемки реже, | 75 |
|  | как сеновал, | 78 |
|  | месяц мимоезжий, | 71 |
|  | переночевал. | 73 |
|  | за телегой | 70 |
|  | впотьмах. | 72 |
|  | день | 75 |
|  | ночлега | 78 |
|  | сеном в волосах. | 82 |
|  | полдень | 67 |
|  | выси, | 64 |
|  | как облака, | 77 |
|  | водка на анисе, | 73 |
|  | и крепка. | 55 |
|  | Казахстан | 461 |
|  | Белоруссия | 453 |
|  | Киргизия | 454 |
|  | Таджикистан | 464 |
|  | Узбекистан | 468 |
|  | Астана | 472 |
|  | Алматы | 459 |
|  | Актау | 458 |
|  | Актобе | 463 |
|  | Атырау | 459 |
|  | Караганда | 459 |
|  | Костанай | 455 |
|  | Кызылорда | 447 |
|  | Павлодар | 446 |
|  | Петропавловск | 446 |
|  | Семей | 413 |
|  | Тараз | 432 |
|  | Уральск | 437 |
|  | Туркестан | 463 |
|  | Кокшетау | 422 |
|  | Шымкент | 451 |
|  | Россия | 452 |
|  | Украина | 447 |
|  | Абхазия | 458 |
|  | Южная Осетия | 457 |
|  | Молдавия | 458 |
|  | Акмолинская | 466 |
|  | Алматинская | 456 |
|  | Мангистауская | 457 |
|  | Актюбинская | 454 |
|  | Атырауская | 452 |
|  | Карагандинская | 455 |
|  | Костанайская | 443 |
|  | Кызылординская | 425 |
|  | Павлодарская | 426 |
|  | Северо-Казахстанская | 437 |
|  | Восточно-Казахстанская | 443 |
|  | Джамбульская | 444 |
|  | Западно-Казахстанская | 427 |
|  | Туркестанская | 459 |
|  | Усть-Каменогорск | 423 |
|  | Южно-Казахстанская | 412 |
|  | Республика | 64 |
|  | республика | 60 |
|  | Казахстан | 61 |
|  | казахстан | 65 |
|  | РК | 62 |
|  | Область | 68 |
|  | область | 66 |
|  | обл | 66 |
|  | обл. | 64 |
|  | город | 68 |
|  | г. | 2 |
|  | Район | 121 |
|  | район | 14 |
|  | р-н | 56 |
|  | Село | 111 |
|  | село | 21 |
|  | с. | 10 |
|  | Поселок | 69 |
|  | поселок | 61 |
|  | п. | 10 |
|  | Атырауская | 65 |
|  | Алматинская | 65 |
|  | Актюбинская | 64 |
|  | Акмолинская | 63 |
|  | Жамбылская | 64 |
|  | Аягоз | 57 |
|  | Восточно-Казахстанская | 64 |
|  | Западно-Казахстанская | 62 |
|  | Южно-Казахстанская | 62 |
|  | Северо-Казахстанская | 61 |
|  | Карагандинская | 61 |
|  | Павлодарская | 63 |
|  | Кызылординская | 65 |
|  | Мангистауская | 65 |
|  | Костанайская | 66 |
|  | Астана | 66 |
|  | Абай | 68 |
|  | Акколь | 68 |
|  | Аксай | 70 |
|  | Аксу | 67 |
|  | Актау | 65 |
|  | Актобе | 66 |
|  | Алга | 67 |
|  | Алматы | 67 |
|  | Аральск | 65 |
|  | Аркалык | 64 |
|  | Арысь | 64 |
|  | Атбасар | 63 |
|  | Атырау | 62 |
|  | Балхаш | 62 |
|  | Байконур | 63 |
|  | Булаево | 45 |
|  | Державинск | 63 |
|  | Ерейментау | 62 |
|  | Есик | 60 |
|  | Есиль | 67 |
|  | Жанаозен | 66 |
|  | Жанатас | 66 |
|  | Жаркент | 67 |
|  | Жезказган | 65 |
|  | Жем | 65 |
|  | Жетысай | 67 |
|  | Житикара | 66 |
|  | Зайсан | 64 |
|  | Зыряновск | 66 |
|  | Казалинск | 64 |
|  | Кандыагаш | 62 |
|  | Капчагай | 64 |
|  | Караганда | 63 |
|  | Каражал | 61 |
|  | Каратау | 63 |
|  | Каркаралинск | 62 |
|  | Каскелен | 64 |
|  | Кентау | 63 |
|  | Кокшетау | 62 |
|  | Костанай | 62 |
|  | Кульсары | 58 |
|  | Курчатов | 45 |
|  | Кызылорда | 61 |
|  | Ленгер | 63 |
|  | Лисаковск | 64 |
|  | Макинск | 68 |
|  | Мамлютка | 67 |
|  | Павлодар | 65 |
|  | Петропавловск | 64 |
|  | Приозерск | 62 |
|  | Риддер | 60 |
|  | Рудный | 60 |
|  | Сарань | 61 |
|  | Сарканд | 62 |
|  | Сарыагаш | 62 |
|  | Сатпаев | 64 |
|  | Семей | 68 |
|  | Сергеевка | 64 |
|  | Серебрянск | 62 |
|  | Степногорск | 63 |
|  | Степняк | 66 |
|  | Тайынша | 68 |
|  | Талгар | 64 |
|  | Талдыкорган | 63 |
|  | Тараз | 63 |
|  | Текели | 64 |
|  | Темир | 61 |
|  | Темиртау | 43 |
|  | Туркестан | 60 |
|  | Уральск | 61 |
|  | Усть-Каменогорск | 62 |
|  | Ушарал | 63 |
|  | Уштобе | 65 |
|  | Форт-Шевченко | 65 |
|  | Хромтау | 65 |
|  | Шалкар | 66 |
|  | Шар | 67 |
|  | Шардара | 63 |
|  | Шахтинск | 62 |
|  | Шемонаиха | 66 |
|  | Шу | 63 |
|  | Шымкент | 64 |
|  | Щучинск | 63 |
|  | Экибастуз | 60 |
|  | Эмба | 60 |
|  | Новодолинский | 63 |
|  | Айтеке би | 63 |
|  | Аксу | 63 |
|  | Аксу | 59 |
|  | Актас | 63 |
|  | Актау | 61 |
|  | Аршалы | 59 |
|  | Атасу | 58 |
|  | Балпык | 58 |
|  | Белколь | 63 |
|  | Бестобе | 63 |
|  | Ботакара | 64 |
|  | Горняцкий | 64 |
|  | Гульшат | 64 |
|  | Деркул | 65 |
|  | Долинка | 65 |
|  | Жайрем | 64 |
|  | Жезказган | 63 |
|  | Заводской | 62 |
|  | Затобольск | 61 |
|  | Зачаганск | 65 |
|  | Индерборский | 62 |
|  | Карабалык | 61 |
|  | Качар | 63 |
|  | Киевка | 64 |
|  | Конырат | 63 |
|  | Круглоозерное | 60 |
|  | Ленинский | 60 |
|  | Макат | 65 |
|  | Октябрьский | 62 |
|  | Осакаровка | 61 |
|  | Отеген батыра | 60 |
|  | Сарыколь | 59 |
|  | Сарыозек | 55 |
|  | Саяк | 52 |

База данных предоставляется бесплатно для научно-исследовательских целей при условии подписания Соглашения о нераспространении и некоммерческом использовании данных.

*Инструкция по получению базы данных:* Набор данных можно скачать по следующей ссылке: <https://github.com/abdoelsayed2016/HKR_Dataset>, заполнив [Application Form](https://github.com/abdoelsayed2016/HKR_Dataset/blob/master/Application_Form/Application_Form_for_HKR.doc) (Приложение 5, Рисунок 5.1) и отправив ее по электронной почте: ([nurseitovdb@gmail.com](mailto:nurseitovdb@gmail.com), [kairat.boss@gmail.com](mailto:kairat.boss@gmail.com)).

По итогам рассмотрения заявки вам отправим пароль от архива на вашу электронную почту.

**Язык программирования**

Python — высокоуровневый язык программирования общего назначения, ориентированный на повышение производительности разработчика и читаемости кода. Синтаксис ядра Python минималистичен. В то же время стандартная библиотека включает большой объём полезных функций. Python поддерживает структурное, объектно-ориентированное, функциональное, императивное и аспектно-ориентированное программирование. Основные архитектурные черты — динамическая типизация, автоматическое управление памятью, полная интроспекция, механизм обработки исключений, поддержка многопоточных вычислений, высокоуровневые структуры данных. Поддерживается разбиение программ на модули, которые, в свою очередь, могут объединяться в пакеты.

Python — активно развивающийся язык программирования, новые версии с добавлением/изменением языковых свойств выходят примерно раз в два с половиной года. Python поддерживает динамическую типизацию, то есть тип переменной определяется только во время исполнения. Поэтому вместо «присваивания значения переменной» лучше говорить о «связывании значения с некоторым именем».

**Тип реализующей ЭВМ**

**Рекомендуемые требования:**

Персональный компьютер с процессором Intel Core i5, ОЗУ – не менее 8 ГБ, объем свободного места на жестком диске – не меньше 20 ГБ, операционная система - Microsoft Windows 10.

Программное обеспечение: архиватор WinRAR не ниже версии 5.0, любая программа для просмотра изображений

**Минимальные требования:**

Персональный компьютер с процессором Intel Pentium IV, ОЗУ – не менее 2 ГБ, объем свободного места на жестком диске – не меньше 6 ГБ, операционная система - Microsoft Windows XP.

Программное обеспечение: архиватор WinRAR не ниже версии 5.0, любая программа для просмотра изображений

**Установка и запуск программы**

Для работы программы анализа базы данных, необходимо установить язык программирования Python 3 и библиотеки numpy и matplotlib. Для установки библиотеки необходимо набрать команду pip install numpy, matplotlib.

Перед запуском программы, необходимо открыть файл dataset\_analyzer.py в любом текстовом редакторе и указать путь к папке с аннотациями в строке номер 9. Затем в командной строке набираем команду python dataset\_analyzer.py и программа начнет расчет. После завершения на экране появиться окно с графиком распределения букв в базе данных (Рисунок 6).

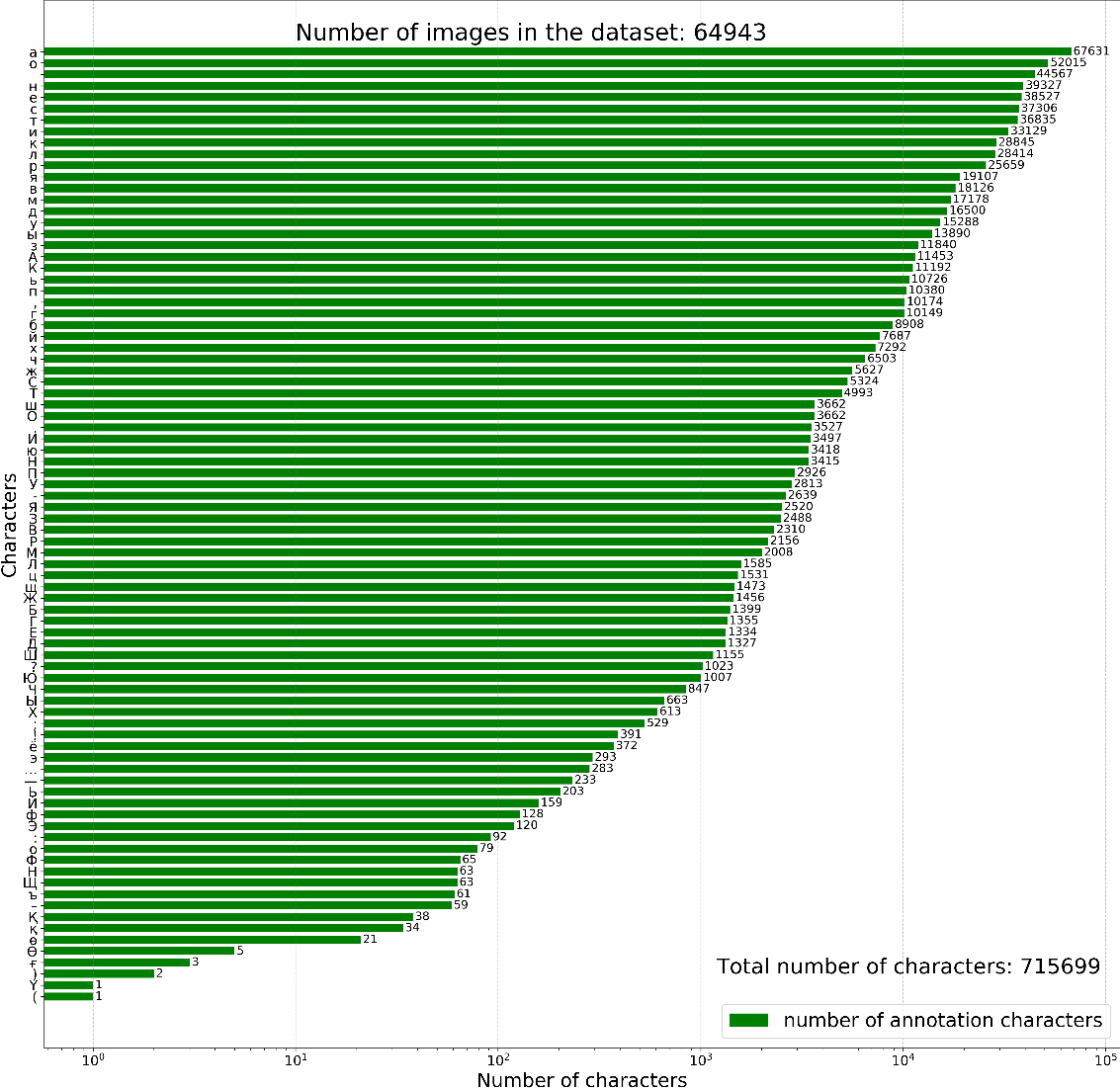
****

Рисунок 6 – График распределения букв в базе данных

Таблица 3 – Табличное представление распределения букв в базе данных

| № | Буква | Количество  символов |
| --- | --- | --- |
|  | а | 67631 |
|  | ә | - |
|  | б | 8908 |
|  | в | 18126 |
|  | г | 10149 |
|  | ғ | 3 |
|  | д | 16500 |
|  | е | 38527 |
|  | ё | 372 |
|  | ж | 5627 |
|  | з | 11840 |
|  | и | 33129 |
|  | й | 7687 |
|  | к | 28845 |
|  | қ | 34 |
|  | л | 28414 |
|  | м | 17178 |
|  | н | 39327 |
|  | ң | - |
|  | о | 52094 |
|  | ө | 21 |
|  | п | 10380 |
|  | р | 25659 |
|  | с | 37306 |
|  | т | 36835 |
|  | у | 15288 |
|  | ұ | - |
|  | ү | - |
|  | ф | 128 |
|  | х | 7292 |
|  | һ | - |
|  | ц | 1531 |
|  | ч | 6503 |
|  | ш | 3662 |
|  | щ | 1473 |
|  | ъ | 61 |
|  | ы | 13890 |
|  | і | - |
|  | ь | 10726 |
|  | э | 293 |
|  | ю | 3418 |
|  | я | 19107 |
|  | А | 11453 |
|  | Ә | - |
|  | Б | 1399 |
|  | В | 2310 |
|  | Г | 1355 |
|  | Ғ | - |
|  | Д | 1327 |
|  | Е | 1334 |
|  | Ё | - |
|  | Ж | 1456 |
|  | З | 2488 |
|  | И | 3497 |
|  | Й | 159 |
|  | К | 11192 |
|  | Қ | 38 |
|  | Л | 1585 |
|  | М | 2008 |
|  | Н | 3478 |
|  | Ң | - |
|  | О | 3662 |
|  | Ө | 5 |
|  | П | 2926 |
|  | Р | 2156 |
|  | С | 5324 |
|  | Т | 4993 |
|  | У | 2813 |
|  | Ұ | - |
|  | Ү | 1 |
|  | Ф | 65 |
|  | Х | 613 |
|  | Һ | - |
|  | Ц | - |
|  | Ч | 847 |
|  | Ш | 1155 |
|  | Щ | 63 |
|  | Ъ | - |
|  | І | - |
|  | Ы | 663 |
|  | Ь | 203 |
|  | Э | 120 |
|  | Ю | 1007 |
|  | Я | 2520 |
|  | пробел | 44567 |
|  | - | 2639 |
|  | – | 59 |
|  | — | 233 |
|  | : | 92 |
|  | ; | 529 |
|  | ! | 391 |
|  | ? | 1023 |
|  | , | 10174 |
|  | . | 3527 |
|  | … | 283 |
|  | ( | 1 |
|  | ) | 2 |

Приложение 1



Рисунок 1.1 – Пример формы первого сбора слов, форма 1



Рисунок 1.2 – Пример формы первого сбора слов, форма 2



Рисунок 1.3 – Пример формы первого сбора слов, форма 3

Приложение 2

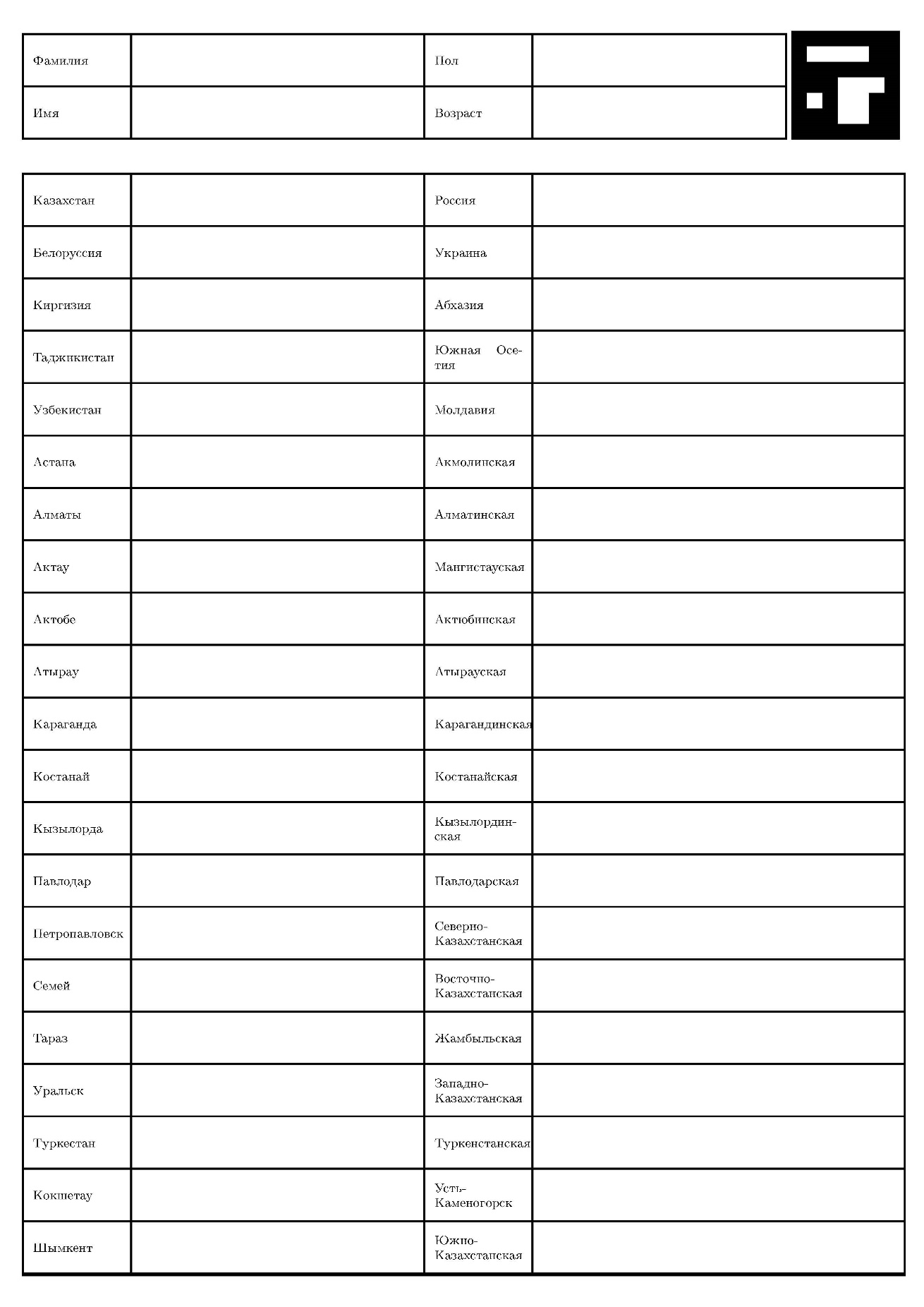


Рисунок 2.1 – Пример формы второго сбора слов

Приложение 3

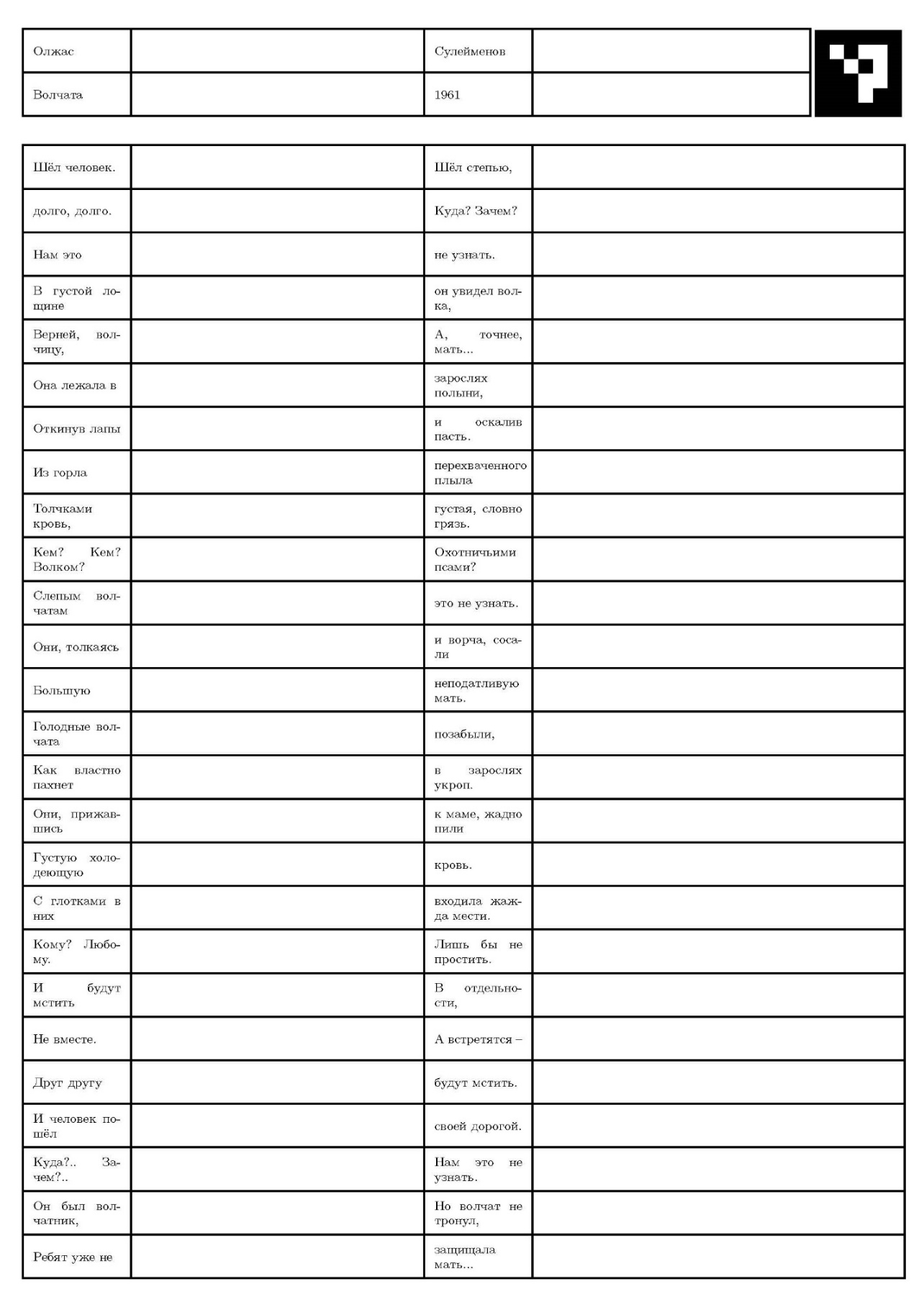


Рисунок 3.1 – Макет формы для заполнения: О. Сулейменов «Волчата», 1961 г.



Рисунок 3.2 – Макет формы для заполнения: О. Сулейменов «Догони! (Кыз куу)», 1961 г.



Рисунок 3.3 – Макет формы для заполнения: М. Лермонтов «Исповедь», 1831 г.

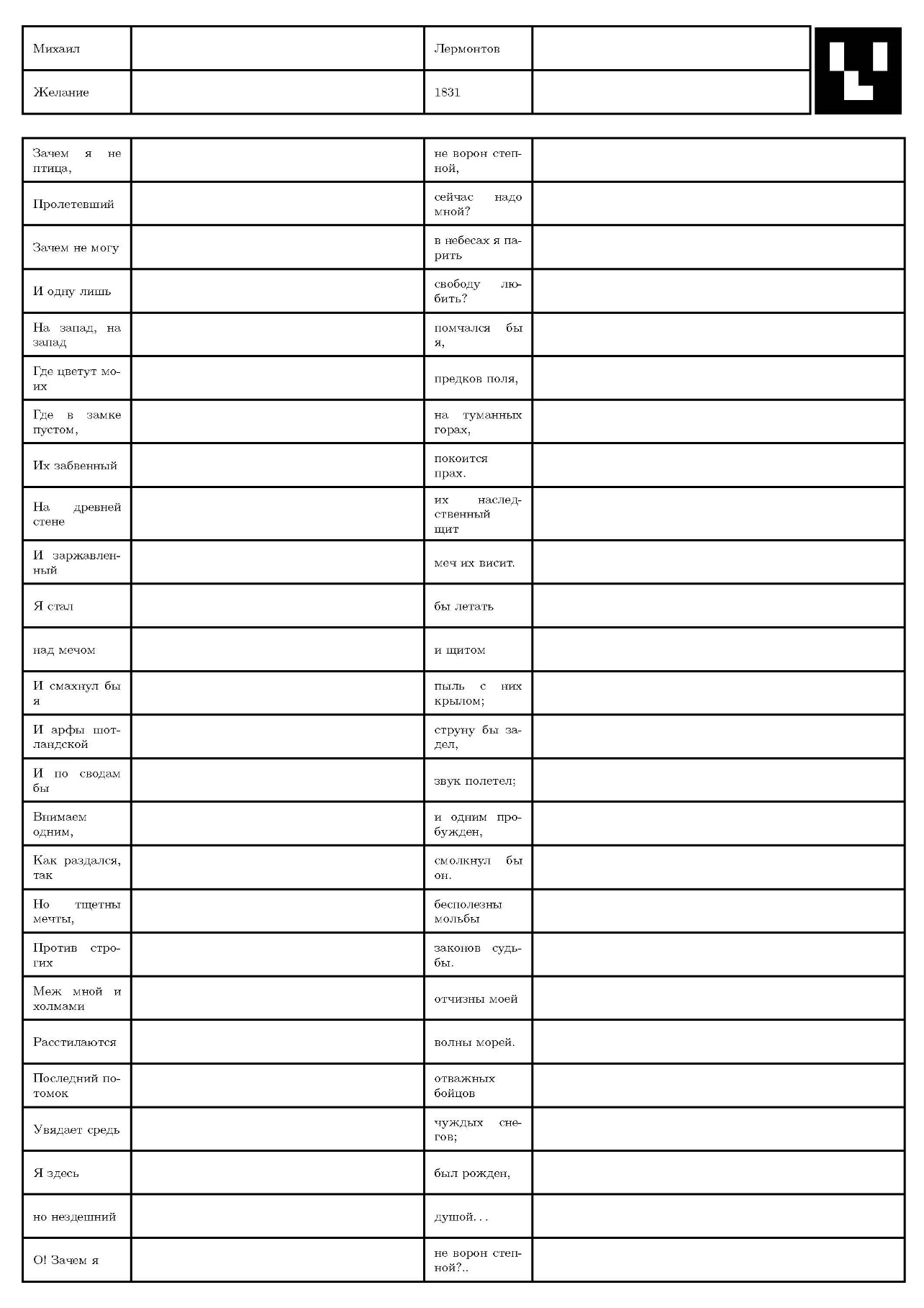


Рисунок 3.4 – Макет формы для заполнения: М. Лермонтов «Желание», 1831 г.



Рисунок 3.5 – Макет формы для заполнения: А. Пушкин «Зимняя дорога», 1826 г.



Рисунок 3.6 – Макет формы для заполнения: А. Ахматова «Прошло пять лет», 1950 г.

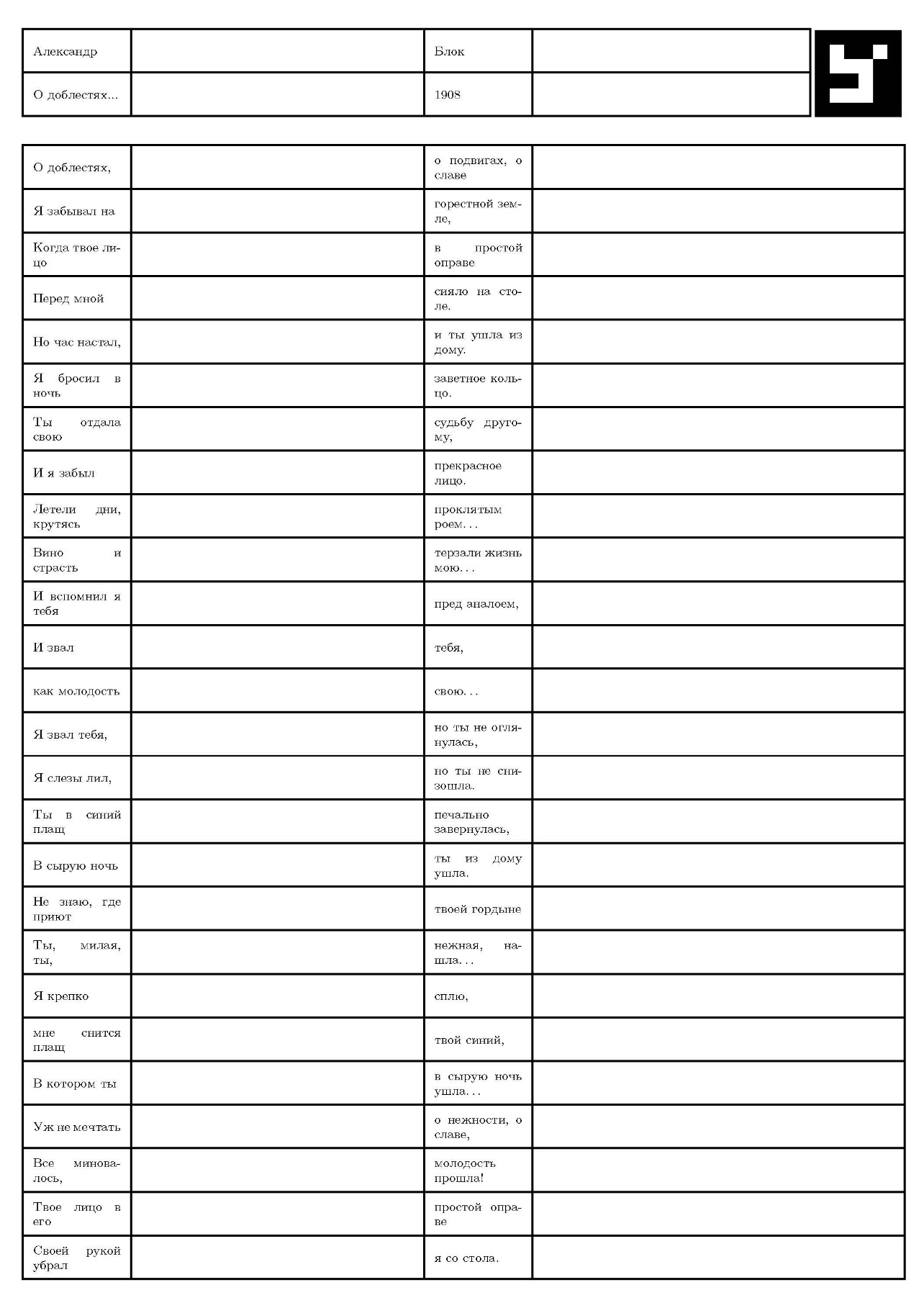


Рисунок 3.7 – Макет формы для заполнения: А. Блок «О доблестях…», 1908 г.

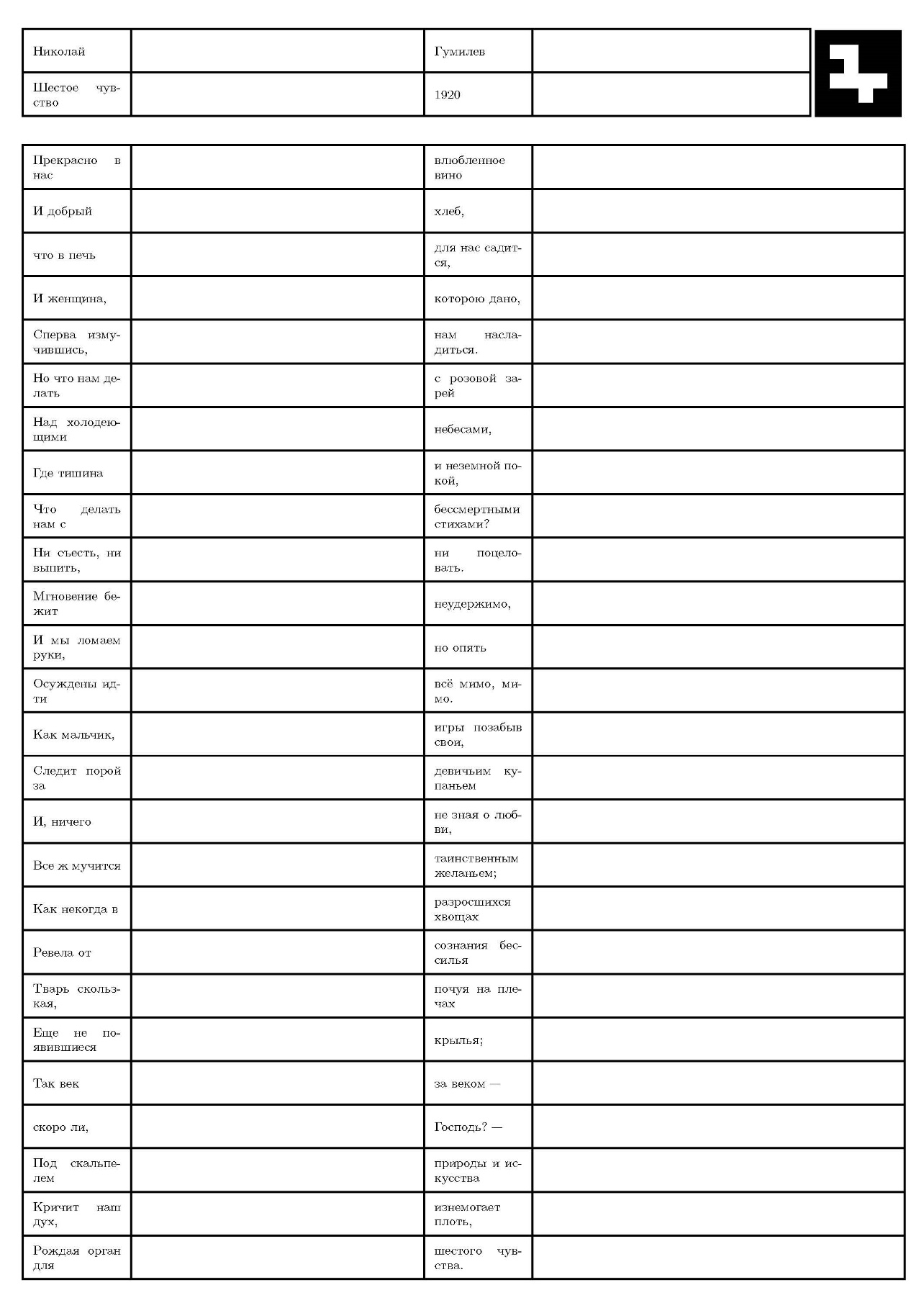


Рисунок 3.8 – Макет формы для заполнения: Н. Гумилев «Шестое чувство», 1920 г.



Рисунок 3.9 – Макет формы для заполнения: И. Бродский «Стих», 1980 г.



Рисунок 3.10 – Макет формы для заполнения: Б. Пастернак «Стога», 1957 г.

Приложение 5

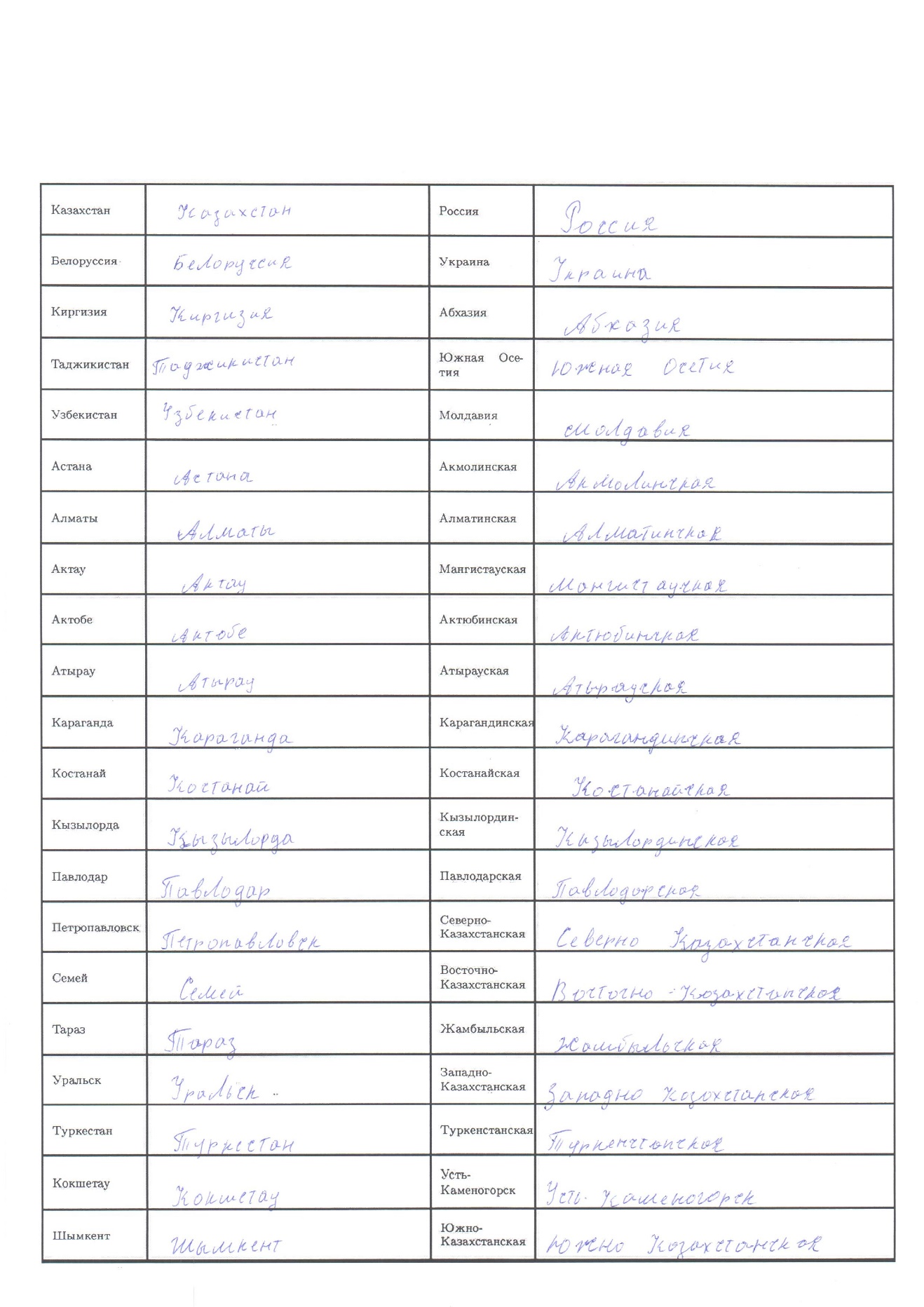


Рисунок 5.1 – Образец заполненной формы

Приложение 6



Рисунок 6.1 – Образец заявки на получение набора данных