

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ ИМЕНИ Н. Э. БАУМАНА  
Факультет информатики и систем управления  
Кафедра теоретической информатики и компьютерных технологий

Наброски дипломного проекта

«Автоматическое установление связей между сообщениями твиттера и  
новостными статьями»

Выполнил:  
студент ИУ9-101  
Выборнов А. И.  
Руководитель:  
Лукашевич Н.В.

Москва 2016

# Содержание

<b>1. Техничко-экономическое обоснование</b>	<b>3</b>
1.1. Трудоемкость разработки программной продукции . . . . .	5
1.1.1. Трудоемкость разработки технического задания . . . . .	5
1.1.2. Трудоемкость разработки эскизного проекта . . . . .	6
1.1.3. Трудоемкость разработки технического проекта . . . . .	7
1.1.4. Трудоемкость разработки рабочего проекта . . . . .	9
1.1.5. Трудоемкость выполнения стадии «Внедрение» . . . . .	11
1.2. Расчет количества исполнителей . . . . .	12
1.3. Ленточный график выполнения работ . . . . .	13
1.4. Определение себестоимости программной продукции . . . . .	14
1.5. Определение стоимости программной продукции . . . . .	15
1.6. Расчет экономической эффективности . . . . .	15
1.7. Вывод . . . . .	16
<b>Список литературы</b>	<b>18</b>

# 1. Техничко-экономическое обоснование

Разработка программного обеспечения — достаточно трудоемкий и длительный процесс, требующий выполнения большого числа разнообразных операций. Организация и планирование процесса разработки программного продукта или программного комплекса при традиционном методе планирования предусматривает выполнение следующих работ:

- формирование состава выполняемых работ и группировка их по стадиям разработки;
- расчет трудоемкости выполнения работ;
- установление профессионального состава и расчет количества исполнителей;
- определение продолжительности выполнения отдельных этапов разработки;
- построение календарного графика выполнения разработки;
- контроль выполнения календарного графика.

Далее приведен перечень и состав работ при разработке программного средства для автоматического установления связей между сообщениями твиттера и новостными статьями. Отметим, что процесс разработки программного продукта характеризуется совместной работой разработчиков постановки задач и разработчиков программного обеспечения.

Укрупненный состав работ по стадиям разработки программного продукта:

## 1. Техническое задание:

- Постановка задач, выбор критериев эффективности,
- Разработка технико-экономического обоснования разработки,
- Определение состава пакета прикладных программ, состава и структуры информационной базы,
- Выбор языков программирования,
- Предварительный выбор методов выполнения работы,
- Разработка календарного плана выполнения работ;

## 2. Эскизный проект:

- Предварительная разработка структуры входных и выходных данных,

- Разработка общего описания алгоритмов реализации решения задач,
- Разработка пояснительной записки,
- Консультации разработчиков постановки задач,
- Согласование и утверждение эскизного проекта;

### 3. Технический проект:

- Разработка алгоритмов решения задач,
- Разработка пояснительной записки,
- Согласование и утверждение технического проекта,
- Разработка структуры программы,
- Разработка программной документации и передача ее для включения в технический проект,
- Уточнение структуры, анализ и определение формы представления входных и выходных данных,
- Выбор конфигурации технических средств;

### 4. Рабочий проект:

- Комплексная отладка задач и сдача в опытную эксплуатацию,
- Разработка проектной документации,
- Программирование и отладка программ,
- Описание контрольного примера,
- Разработка программной документации,
- Разработка, согласование программы и методики испытаний,
- Предварительное проведение всех видов испытаний;

### 5. Внедрение:

- Подготовка и передача программной документации для сопровождения с оформлением соответствующего Акта,
- Передача программной продукции в фонд алгоритмов и программ,
- Проверка алгоритмов и программ решения задач, корректировка документации после опытной эксплуатации программного продукта;

Трудоемкость разработки программной продукции зависит от ряда факторов, основными из которых являются следующие: степень новизны разрабатываемого программного комплекса, сложность алгоритма его функционирования, объем используемой информации, вид ее представления и способ обработки, а также уровень используемого алгоритмического языка программирования. Чем выше уровень языка, тем трудоемкость меньше.

По степени новизны разрабатываемый проект относится к *группе новизны А* – разработка программных комплексов, требующих использования принципиально новых методов их создания, проведения НИР и т.п.

По степени сложности алгоритма функционирования проект относится к *2 группе сложности* - программная продукция, реализующая учетно-статистические алгоритмы.

По виду представления исходной информации и способа ее контроля программный продукт относится к *группе 12* - исходная информация представлена в форме документов, имеющих различный формат и структуру и *группе 22* - требуется печать документов одинаковой формы и содержания, вывод массивов данных на машинные носители.

## 1.1. Трудоемкость разработки программной продукции

Трудоемкость разработки программной продукции ( $\tau_{PP}$ ) может быть определена как сумма величин трудоемкости выполнения отдельных стадий разработки программного продукта из выражения:

$$\tau_{PP} = \tau_{TZ} + \tau_{EP} + \tau_{TP} + \tau_{RP} + \tau_V,$$

где  $\tau_{TZ}$  — трудоемкость разработки технического задания на создание программного продукта;  $\tau_{EP}$  — трудоемкость разработки эскизного проекта программного продукта;  $\tau_{TP}$  — трудоемкость разработки технического проекта программного продукта;  $\tau_{RP}$  — трудоемкость разработки рабочего проекта программного продукта;  $\tau_V$  — трудоемкость внедрения разработанного программного продукта.

### 1.1.1. Трудоемкость разработки технического задания

Расчёт трудоёмкости разработки технического задания ( $\tau_{PP}$ ) [чел.-дни] производится по формуле:

$$\tau_{TZ} = T_{RZ}^Z + T_{RP}^Z,$$

где  $T_{RZ}^Z$  — затраты времени разработчика постановки задачи на разработку ТЗ, [чел.-дни];  $T_{RP}^Z$  — затраты времени разработчика программного обеспечения на разработку ТЗ, [чел.-дни]. Их значения рассчитываются по формулам:

$$T_{RZ}^Z = t_Z * K_{RZ}^Z,$$

$$T_{RP}^Z = t_Z * K_{RP}^Z,$$

где  $t_Z$  — норма времени на разработку ТЗ на программный продукт (зависит от функционального назначения и степени новизны разрабатываемого программного продукта), [чел.-дни]. В нашем случае по таблице получаем значение (группа новизны — А, функциональное назначение — технико-экономическое планирование):

$$t_Z = 79.$$

$K_{RZ}^Z$  — коэффициент, учитывающий удельный вес трудоемкости работ, выполняемых разработчиком постановки задачи на стадии ТЗ. В нашем случае (совместная разработка с разработчиком ПО):

$$K_{RZ}^Z = 0.65.$$

$K_{RP}^Z$  — коэффициент, учитывающий удельный вес трудоемкости работ, выполняемых разработчиком программного обеспечения на стадии ТЗ. В нашем случае (совместная разработка с разработчиком постановки задач):

$$K_{RP}^Z = 0.35.$$

Тогда:

$$\tau_{TZ} = 79 * (0.35 + 0.65) = 79.$$

### 1.1.2. Трудоемкость разработки эскизного проекта

Расчёт трудоёмкости разработки эскизного проекта ( $\tau_{EP}$ ) [чел.-дни] производится по формуле:

$$\tau_{EP} = T_{RZ}^E + T_{RP}^E,$$

где  $T_{RZ}^E$  — затраты времени разработчика постановки задачи на разработку эскизного проекта (ЭП), [чел.-дни];  $T_{RP}^E$  — затраты времени разработчика программного обеспечения на разработку ЭП, [чел.-дни]. Их значения рассчитываются по форму-

лам:

$$T_{RZ}^E = t_E * K_{RZ}^E,$$

$$T_{RP}^E = t_E * K_{RP}^E,$$

где  $t_E$  – норма времени на разработку ЭП на программный продукт (зависит от функционального назначения и степени новизны разрабатываемого программного продукта), [чел.-дни]. В нашем случае по таблице получаем значение (группа новизны – А, функциональное назначение – технико-экономическое планирование):

$$t_E = 175.$$

$K_{RZ}^E$  – коэффициент, учитывающий удельный вес трудоемкости работ, выполняемых разработчиком постановки задачи на стадии ЭП. В нашем случае (совместная разработка с разработчиком ПО):

$$K_{RZ}^E = 0.7.$$

$K_{RP}^E$  – коэффициент, учитывающий удельный вес трудоемкости работ, выполняемых разработчиком программного обеспечения на стадии ТЗ. В нашем случае (совместная разработка с разработчиком постановки задач):

$$K_{RP}^E = 0.3.$$

Тогда:

$$\tau_{EP} = 175 * (0.3 + 0.7) = 175.$$

### 1.1.3. Трудоемкость разработки технического проекта

Трудоёмкость разработки технического проекта ( $\tau_{TP}$ ) [чел.-дни] зависит от функционального назначения программного продукта, количества разновидностей форм входной и выходной информации и определяется по формуле:

$$\tau_{TP} = (t_{RZ}^T + t_{RP}^T) * K_V * K_R,$$

где  $t_{RZ}^T$  – норма времени, затрачиваемого на разработку технического проекта (ТП) разработчиком постановки задач, [чел.-дни];  $t_{RP}^T$  – норма времени, затрачиваемого на разработку ТП разработчиком ПО, [чел.-дни]. По таблице принимаем (функциональное назначение — технико-экономическое планирование, количество разновид-

ностей форм входной информации — 2 (твиты, новости), количество разновидностей форм выходной информации — 2 (набор связей твит-новости, оценка работы рекомендательной системы)):

$$t_{RZ}^T = 52,$$

$$t_{RP}^T = 14.$$

$K_R$  — коэффициент учета режима обработки информации. По таблице принимаем (группа новизны — А, режим обработки информации — реальный масштаб времени):

$$K_R = 1.67.$$

$K_V$  — коэффициент учета вида используемой информации, определяется по формуле:

$$K_V = \frac{K_P * n_P + K_{NS} * n_{NS} + K_B * n_B}{n_P + n_{NS} + n_B},$$

где  $K_P$  — коэффициент учета вида используемой информации для переменной информации;  $K_{NS}$  — коэффициент учета вида используемой информации для нормативно-справочной информации;  $K_B$  — коэффициент учета вида используемой информации для баз данных;  $n_P$  — количество наборов данных переменной информации;  $n_{NS}$  — количество наборов данных нормативно-справочной информации;  $n_B$  — количество баз данных. Коэффициенты находим по таблице (группа новизны - А):

$$K_P = 1.70,$$

$$K_{NS} = 1.45,$$

$$K_B = 4.37.$$

Количество наборов данных, используемых в рамках задачи:

$$n_P = 3,$$

$$n_{NS} = 0,$$

$$n_B = 1.$$

Находим значение  $K_V$ :

$$K_V = \frac{1.70 * 3 + 1.45 * 0 + 4.37 * 1}{3 + 0 + 1} = 2.3675.$$



Тогда:

$$\tau_{TP} = (52 + 14) * 2.3675 * 1.67 = 261.$$

#### 1.1.4. Трудоемкость разработки рабочего проекта

Трудоёмкость разработки рабочего проекта ( $\tau_{RP}$ ) [чел.-дни] зависит от функционального назначения программного продукта, количества разновидностей форм входной и выходной информации, сложности алгоритма функционирования, сложности контроля информации, степени использования готовых программных модулей, уровня алгоритмического языка программирования и определяется по формуле:

$$\tau_{RP} = (t_{RZ}^R + t_{RP}^R) * K_K * K_R * K_Y * K_Z * K_{IA},$$

где  $t_{RZ}^R$  — норма времени, затраченного на разработку рабочего проекта на алгоритмическом языке высокого уровня разработчиком постановки задач, [чел.-дни].  $t_{RP}^R$  — норма времени, затраченного на разработку рабочего проекта на алгоритмическом языке высокого уровня разработчиком ПО, [чел.-дни]. По таблице принимаем (функциональное назначение — технико-экономическое планирование, количество разновидностей форм входной информации — 2 (твиты, новости), количество разновидностей форм выходной информации — 2 (набор связей твит-новости, оценка работы рекомендательной системы)):

$$t_{RZ}^R = 15,$$

$$t_{RP}^R = 91.$$

$K_K$  — коэффициент учета сложности контроля информации. По таблице принимаем (степень сложности контроля входной информации — 12, степень сложности контроля выходной информации — 22):

$$K_K = 1.00.$$

$K_R$  — коэффициент учета режима обработки информации. По таблице принимаем (группа новизны — А, режим обработки информации — реальный масштаб времени):

$$K_R = 1.75.$$

$K_Y$  — коэффициент учета уровня используемого алгоритмического языка программирования. По таблице принимаем значение (интерпретаторы, языковые описатели):

$$K_Y = 0.8.$$

$K_Z$  — коэффициент учета степени использования готовых программных модулей. По таблице принимаем (использование готовых программных модулей составляет около 30

$$K_Z = 0.7.$$

$K_{IA}$  — коэффициент учета вида используемой информации и сложности алгоритма программного продукта, его значение определяется по формуле:

$$K_{IA} = \frac{K'_P * n_P + K'_{NS} * n_{NS} + K'_B * n_B}{n_P + n_{NS} + n_B},$$

где  $K'_P$  — коэффициент учета сложности алгоритма ПП и вида используемой информации для переменной информации;  $K'_{NS}$  — коэффициент учета сложности алгоритма ПП и вида используемой информации для нормативно-справочной информации;  $K'_B$  — коэффициент учета сложности алгоритма ПП и вида используемой информации для баз данных.  $n_P$  — количество наборов данных переменной информации;  $n_{NS}$  — количество наборов данных нормативно-справочной информации;  $n_B$  — количество баз данных. Коэффициенты находим по таблице (группа новизны - А):

$$K'_P = 2.02,$$

$$K'_{NS} = 1.21,$$

$$K'_B = 1.05.$$

Количество наборов данных, используемых в рамках задачи:

$$n_P = 3,$$

$$n_{NS} = 0,$$

$$n_B = 1.$$

Находим значение  $K_{IA}$ :

$$K_{IA} = \frac{2.02 * 3 + 1.21 * 0 + 1.05 * 1}{3 + 0 + 1} = 1.7775.$$

Тогда:

$$\tau_{RP} = (15 + 91) * 1.00 * 1.75 * 0.8 * .7 * 1.7775 = 185.$$

### 1.1.5. Трудоемкость выполнения стадии «Внедрение»

Расчёт трудоёмкости разработки технического проекта ( $\tau_V$ ) [чел.-дни] производится по формуле:

$$\tau_V = (t_{RZ}^V + t_{RP}^V) * K_K * K_R * K_Z,$$

где  $t_{RZ}^V$  — норма времени, затрачиваемого разработчиком постановки задач на выполнение процедур внедрения программного продукта, [чел.-дни];  $t_{RP}^V$  — норма времени, затрачиваемого разработчиком программного обеспечения на выполнение процедур внедрения программного продукта, [чел.-дни]. По таблице принимаем (функциональное назначение — технико-экономическое планирование, количество разновидностей форм входной информации — 2 (твиты, новости), количество разновидностей форм выходной информации — 2 (набор связей твит-новости, оценка работы рекомендательной системы)):

$$t_{RZ}^V = 17,$$

$$t_{RP}^V = 19.$$

Коэффициент  $K_K$  и  $K_Z$  были найдены выше:

$$K_K = 1.00,$$

$$K_Z = 0.7.$$

$K_R$  — коэффициент учета режима обработки информации. По таблице принимаем (группа новизны — А, режим обработки информации — реальный масштаб времени):

$$K_R = 1.60.$$

Тогда:

$$\tau_V = (17 + 19) * 1.00 * 1.60 * 0.7 = 40.$$

Общая трудоёмкость разработки ПП:

$$\tau_{RP} = 79 + 175 + 261 + 185 + 40 = 740.$$

## 1.2. Расчет количества исполнителей

Средняя численность исполнителей при реализации проекта разработки и внедрения ПО определяется соотношением:

$$N = \frac{t}{F},$$

где  $t$  — затраты труда на выполнение проекта (разработка и внедрение ПО);  $F$  — фонд рабочего времени. Разработка велась 5 месяцев с 1 января 2016 по 31 мая 2016. Количество рабочих дней по месяцам приведено в таблице 1. Из таблицы получаем, что фонд рабочего времени

$$F = 96.$$

Таблица 1: Количество рабочих дней по месяцам

Номер месяца	Интервал дней	Количество рабочих дней
1	01.01.2016 - 31.01.2016	15
3	01.02.2016 - 29.02.2016	20
4	01.03.2016 - 31.03.2016	21
5	01.04.2016 - 30.04.2016	21
6	01.05.2016 - 31.05.2016	19
Итого		96

Получаем число исполнителей проекта:

$$N = \frac{740}{96} = 8$$

Для реализации проекта потребуются 3 старших инженеров и 5 простых инженеров.

### 1.3. Ленточный график выполнения работ

На основе рассчитанных в главах 1.1, 1.2 трудоёмкости и фонда рабочего времени найдём количество рабочих дней, требуемых для выполнения каждого этапа разработка. Результаты приведены в таблице 2.

Таблица 2: Трудоёмкость выполнения работы над проектом

Номер стадии	Название стадии	Трудоёмкость [чел.-дни]	Удельный вес [%]	Количество рабочих дней
1	Техническое задание	79	11	10
2	Эскизный проект	175	24	23
3	Технический проект	261	35	34
4	Рабочий проект	185	25	24
5	Внедрение	40	5	5
Итого		740	100	96

Планирование и контроль хода выполнения разработки проводится по ленточному графику выполнения работ. По данным в таблице 2 в ленточный график (таблица 3), в ячейки столбца “продолжительности рабочих дней” заносятся времена, которые требуются на выполнение соответствующего этапа. Все исполнители работают одновременно.

Таблица 3: Ленточный график выполнения работ

Номер стадии		Продолжительность [раб.-дни]	Календарные дни																							
			Количество рабочих дней																							
			01.01.2016 - 03.01.2016	04.01.2016 - 10.01.2016	11.01.2016 - 17.01.2016	18.01.2016 - 24.01.2016	25.01.2016 - 31.01.2016	01.02.2016 - 07.02.2016	08.02.2016 - 14.02.2016	15.02.2016 - 21.02.2016	22.02.2016 - 28.02.2016	29.02.2016 - 06.03.2016	07.03.2016 - 13.03.2016	14.03.2016 - 20.03.2016	21.03.2016 - 27.03.2016	28.03.2016 - 03.04.2016	04.04.2016 - 10.04.2016	11.04.2016 - 17.04.2016	18.04.2016 - 24.04.2016	25.04.2016 - 01.05.2016	02.05.2016 - 08.05.2016	08.05.2016 - 15.05.2016	16.05.2016 - 22.05.2016	23.05.2016 - 29.05.2016	30.05.2016 - 31.05.2016	
1	10			5	5		5		5		3	5	3	5	5	5	5	5	5	3	4	5	5	2		
2	23					5	5	5	6	2																
3	34									1	5	3	5	5	5	5	5									
4	24																	5	5	3	4	5	2			
5	5																					3	2			

## 1.4. Определение себестоимости программной продукции

Затраты, образующие себестоимость продукции (работ, услуг), состоят из затрат на заработную плату исполнителям, затрат на закупку или аренду оборудования, затрат на организацию рабочих мест, и затрат на накладные расходы.

В таблице 4 приведены затраты на заработную плату и отчисления на социальное страхование в пенсионный фонд, фонд занятости и фонд обязательного медицинского страхования (30.5 %). Для старшего инженера предполагается оклад в размере 120000 рублей в месяц, для инженера предполагается оклад в размере 100000 рублей в месяц.

Таблица 4: Затраты на зарплату и отчисления на социальное страхование

Должность	Зарплата в месяц	Рабочих месяцев	Суммарная зарплата	Затраты на социальные нужды
Старший инженер	120000	5	600000	183000
Старший инженер	120000	5	600000	183000
Старший инженер	120000	5	600000	183000
Инженер	100000	5	500000	152500
Инженер	100000	5	500000	152500
Инженер	100000	5	500000	152500
Инженер	100000	5	500000	152500
Инженер	100000	5	500000	152500
Суммарные затраты			5611500	

Расходы на материалы, необходимые для разработки программной продукции, указаны в таблице 5.

Таблица 5: Затраты на материалы

Наименование материала	Единица измерения	Кол-во	Цена за единицу, руб.	Сумма, руб.
Бумага А4	Пачка 400 л.	2	200	400
Картридж для принтера HP P10025	Шт.	3	450	1350
Суммарные затраты				1750

В работе над проектом используется специальное оборудование — персональные электронно-вычислительные машины (ПЭВМ) в количестве 9 шт. Стоимость одной ПЭВМ составляет 90000 рублей. Месячная норма амортизации  $K = 2,7\%$ . Тогда за 4 месяцев работы расходы на амортизацию составят  $P = 90000 * 9 * 0.027 * 4 = 87480$  рублей.

Общие затраты на разработку программного продукта (ПП) составят  
 $5611500 + 1750 + 87480 = 5700730$  рублей.

## 1.5. Определение стоимости программной продукции

Для определения стоимости работ необходимо на основании плановых сроков выполнения работ и численности исполнителей рассчитать общую сумму затрат на разработку программного продукта. Если ПП рассматривается и создается как продукция производственно-технического назначения, допускающая многократное тиражирование и отчуждение от непосредственных разработчиков, то ее цена  $P$  определяется по формуле:

$$P = K * C + Pr,$$

где  $C$  — затраты на разработку ПП (сметная себестоимость);  $K$  — коэффициент учёта затрат на изготовление опытного образца ПП как продукции производственно-технического назначения ( $K = 1.1$ );  $Pr$  — нормативная прибыль, рассчитываемая по формуле:

$$Pr = \frac{C * \rho_N}{100},$$

где  $\rho_N$  — норматив рентабельности,  $\rho_N = 30\%$ ;

Получаем стоимость программного продукта:

$$P = 1.1 * 5700730 + 5700730 * 0.3 = 7981022 \text{ рублей.}$$

## 1.6. Расчет экономической эффективности

Основными показателями экономической эффективности является чистый дисконтированный доход (PDD) и срок окупаемости вложенных средств. Чистый дисконтированный доход определяется по формуле:

$$PDD = \sum_{t=0}^T (R_t - Z_t) * \frac{1}{(1 + E)^t},$$

где  $T$  — горизонт расчета по месяцам;  $t$  — период расчета;  $R_t$  — результат, достигнутый на  $t$  шаге (стоимость);  $Z_t$  — текущие затраты (на шаге  $t$ );  $E$  — приемлемая для инвестора норма прибыли на вложенный капитал.

На момент начала 2016 года, ставка рефинансирования 11% годовых (ЦБ РФ), что эквивалентно (0.87% в месяц). В виду особенности разрабатываемого продукта

он может быть продан лишь однократно. Отсюда получаем

$$E = 0.0087.$$

В таблице 6 находится расчёт чистого дисконтированного дохода. График его изменения приведён на рисунке 1.

Таблица 6: Расчёт чистого дисконтированного дохода

Месяц	Текущие затраты, руб.	Затраты с начала года, руб.	Текущий доход, руб.	ЧДД, руб.
Январь	1211530	1211530	0	-1211530
Февраль	1122300	2333830	0	-2324150
Март	1122300	3456130	0	-3427174
Апрель	1122300	4578430	0	-4520684
Мая	1122300	5700730	7981022	2104456

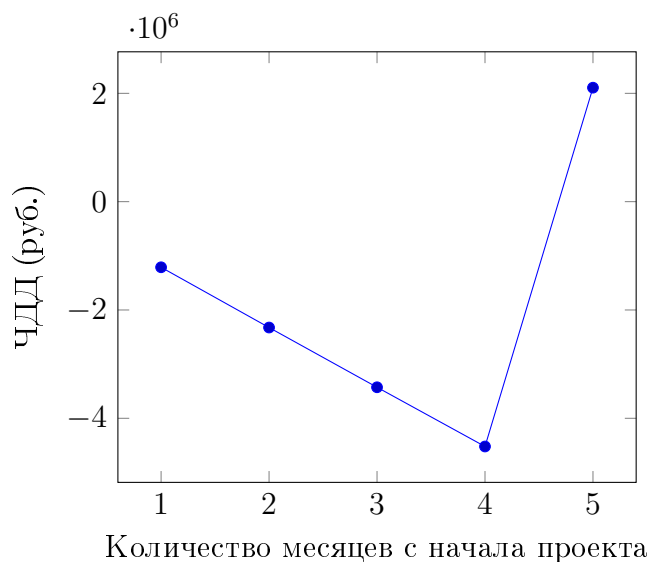


Рисунок 1 — График изменения чистого дисконтированного дохода

Согласно проведенным расчетам, проект является рентабельным. Разрабатываемый проект позволит превысить показатели качества существующих систем и сможет их заменить. Итоговый ЧДД составил: 2104456 рублей.

## 1.7. Результаты

В рамках организационно-экономической части был спланирован календарный график проведения работ по созданию подсистемы поддержки проведения диа-



гностики промышленных, а также были проведены расчеты по трудозатратам. Были исследованы и рассчитаны следующие статьи затрат: материальные затраты; заработная плата исполнителей; отчисления на социальное страхование; накладные расходы.

В результате расчетов было получено общее время выполнения проекта, которое составило 96 рабочих дней, получены данные по суммарным затратам на создание системы для автоматического сопоставления твитов и новостных статей, которые составили 5700730 рублей. Согласно проведенным расчетам, проект является рентабельным. Цена данного программного проекта составила 7981022 рублей, итоговый ЧДД составил 2104456 рублей.

## Список литературы

- [1] W. Guo, H. Li, H. Ji, and M. T. Diab. Linking tweets to news: A framework to enrich short text data in social media. - ACL, pages 239–249, 2013.
- [2] Manos Tsagkias, Maarten de Rijke, Wouter Weerkamp. Linking Online News and Social Media. - ISLA, University of Amsterdam.
- [3] T. Hoang-Vu, A. Bessa, L. Barbosa and J. Freire. Bridging Vocabularies to Link Tweets and News. - International Workshop on the Web and Databases (WebDB 2014), Snowbird, Utah, US, 2014.
- [4] J. Sankaranarayanan, H. Samet, B. Teitler, M. Lieberman, J. Sperling. TwitterStand: news in tweets. - 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 2009, Seattle, Washington.
- [5] Ou Jin, Nathan N. Liu, Kai Zhao, Yong Yu, and Qiang Yang. 2011. Transferring topical knowledge from auxiliary long texts for short text clustering. In Proceedings of the 20th ACM international conference on Information and knowledge management.
- [6] Weiwei Guo and Mona Diab. 2012a. Modeling sentences in the latent space. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics.
- [7] Harald Steck. 2010. Training and testing of recommender systems on data missing not at random. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
- [8] Nathan Srebro and Tommi Jaakkola. 2003. Weighted low-rank approximations. In Proceedings of the Twentieth International Conference on Machine Learning.
- [9] Eric Huns. Hunseblog on Wordpress: URL: <https://hunseblog.wordpress.com/2014/09/15/installing-numpy-and-openblas/>.
- [10] Арсеньев В.В., Сажин Ю.Б. Методические указания к выполнению организационно-экономической части дипломных проектов по созданию программной продукции. М.: изд. МГТУ им. Баумана, 1994. 52 с. 2.
- [11] Под ред. Смирнова С.В. Организационно-экономическая часть дипломных проектов исследовательского профиля. М.: изд. МГТУ им. Баумана, 1995. 100 с.

[12] ГОСТ 34.601 "АС. Стадии создания".

---