

# Использование автоматической классификации изменений программного кода

в управлении процессом  
разработки программного  
обеспечения

***Князев Е.Г.***

Санкт-Петербургский Государственный Университет Информационных Технологий, Механики и Оптики; ЗАО «Транзас Технологии»

***Шопырин Д.Г.***

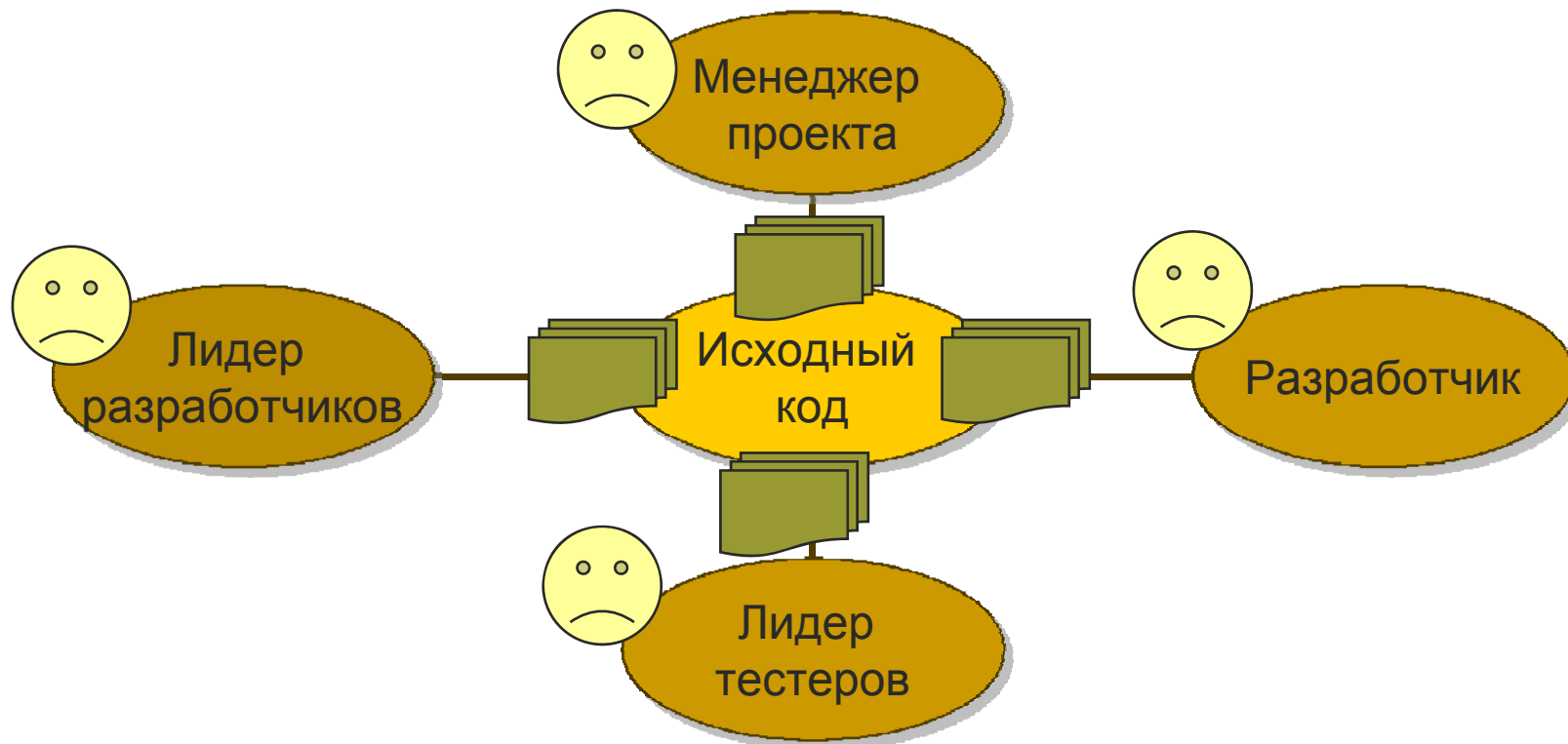
Санкт-Петербургский Государственный Университет Информационных Технологий, Механики и Оптики

# [ Содержание ]

---

- Просмотр исходного кода
- Анализ истории исходного кода
- Автоклассификация изменений
- Будущие исследования

# [ Просмотр исходного кода ]



# [ Разработчик ]

- Просмотр большого количества изменений
- Commit

<i>Проект</i>	<i>Tortoise SVN</i>	<i>Navi- Manager</i>	<i>KDE</i>
<i>Период ~1 мес.</i>	22.09.2007 22.10.2007		17.09.2007 14.10.2007
<i>Кол-во измен.</i>	<b>215</b>	<b>72</b>	<b>11841 (!)</b>

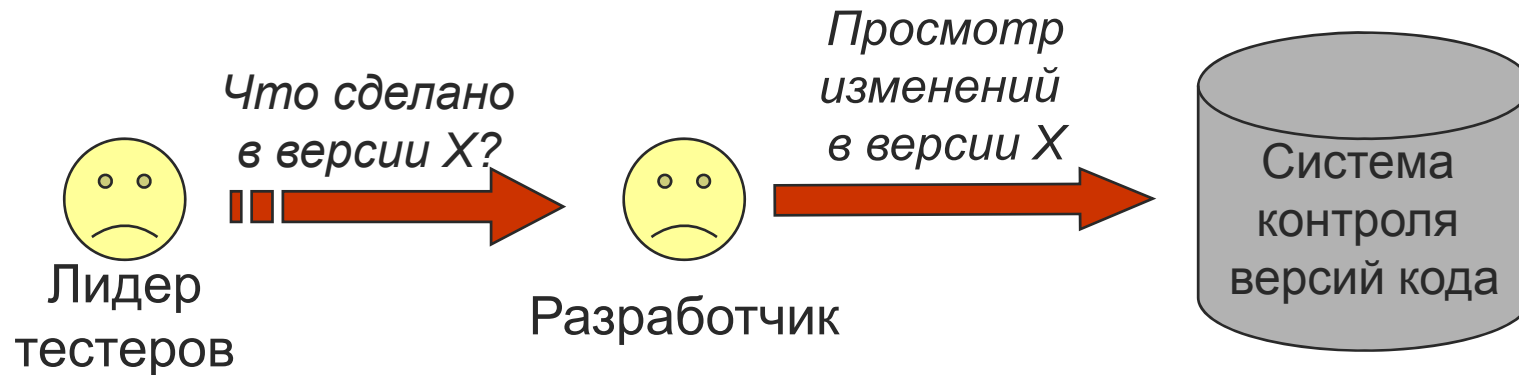
# Лидер разработчиков

- Просмотр большого количества изменений
- Запрет некоторых типов изменений

<i>Стадия разработки</i>	<i>Реализация</i>	<i>Тестирование</i>	<i>Поддержка</i>
<i>Тип изменения</i>			
<i>Новая функциональность</i>	+	—	—
<i>Удаление кода</i>	+	—	—
<i>Рефакторинг</i>	+	—	—
<i>Незначительное исправление</i>	+	+	—
<i>Критичное исправление</i>	+	+	+

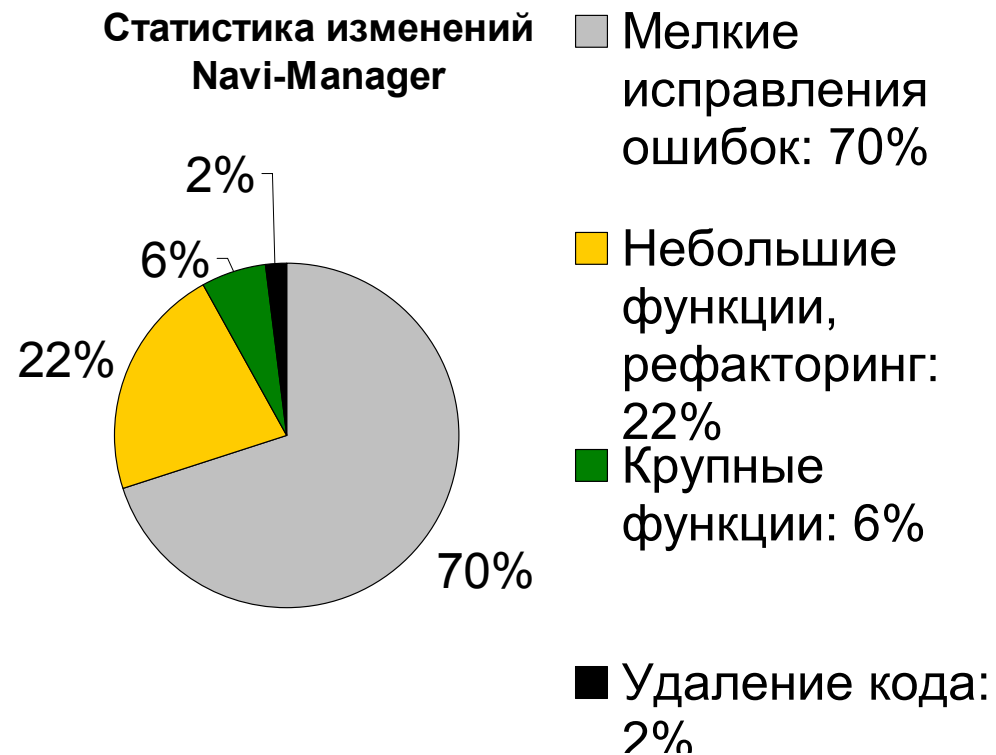
# [ Лидер тестеров ]

- Получение списка функциональности для конкретной версии



# [ Менеджер проекта ]

- Анализ параметров процесса разработки



# [ Метод решения проблемы ]

- **Проблема**

- Большое количество изменений для просмотра

- **Решение**

- Сужение множества просмотра

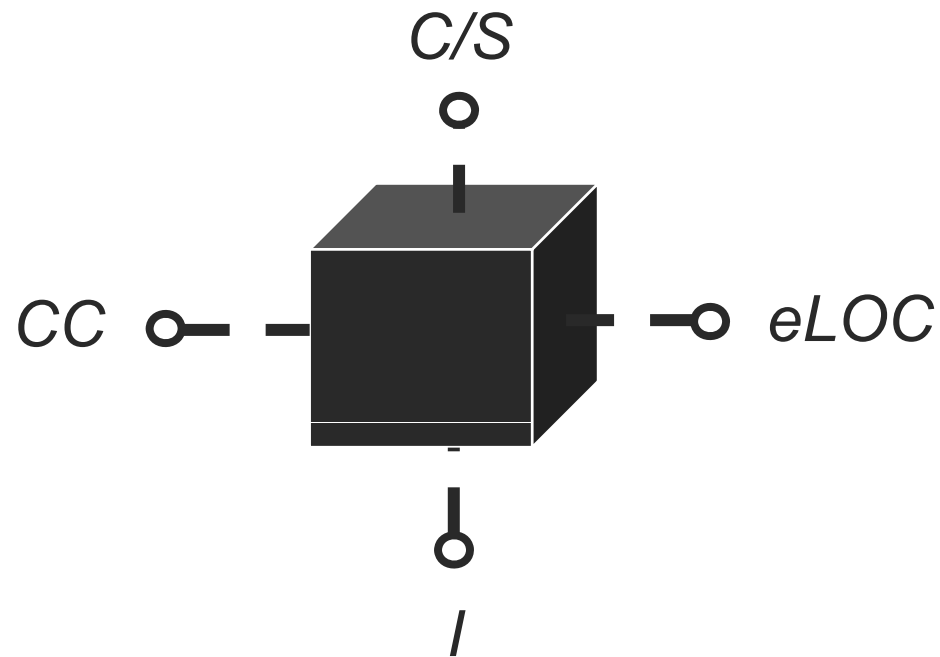
- **Метод**

- Автоматическая классификация изменений



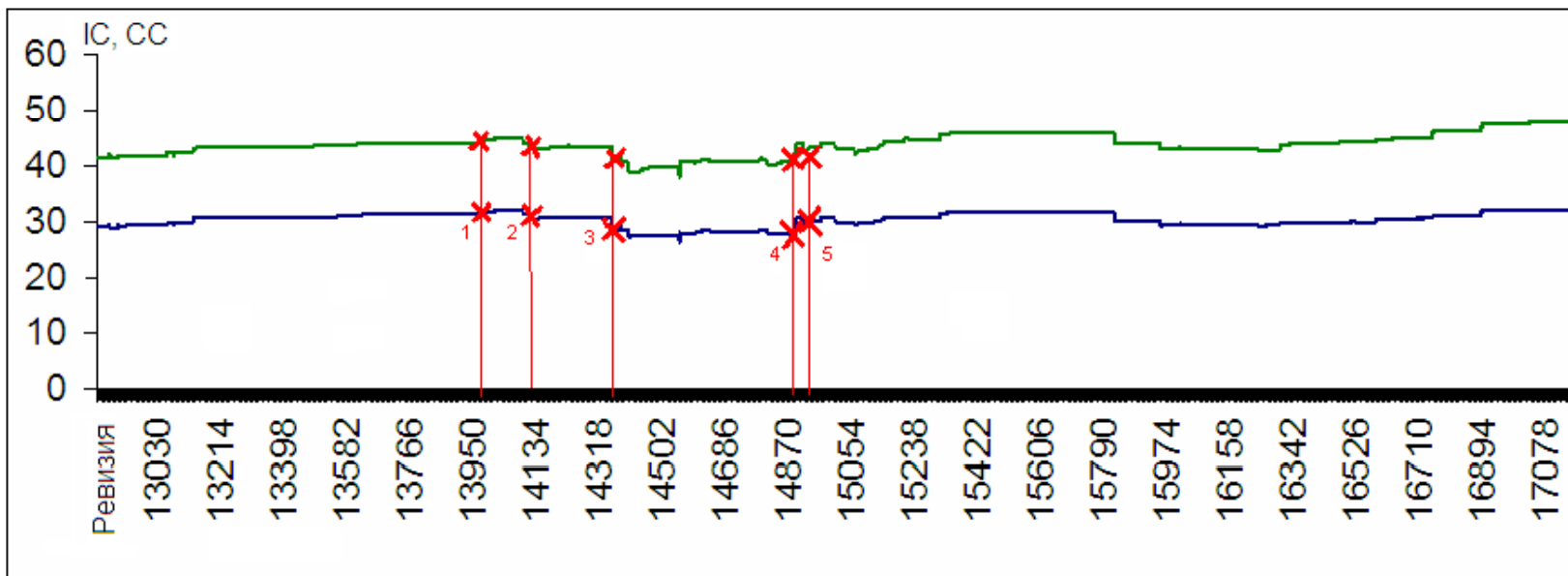
# [ Анализ истории кода ]

- Недостатки статических метрик
- Анализ исторических метрик



# Недостатки статических метрик

- Характеризуют мгновенные состояния



# Недостатки статических метрик

- Сложно интерпретировать значение
  - Нестабильность  $I = C_a / (C_a + C_e)$
  - $C_a$  – внутреннее сцепление
  - $C_e$  – внешнее сцепление

# [ Допустимые значения метрик ]

- Винников М. А., Панекин Н. Н.  
Определение информативности метрик программного кода // *Вестник ДГМА №1 2006, с. 13-18*
  - *CC* – цикломатическая сложность
  - *LOCm* – число строк кода в методе
  - *I* – нестабильность
  - *A* – абстрактность

# Допустимые значения метрик

- Средние значения метрик и разброс для 18 тестовых проектов на Java

<i>Метрика</i>	<i>Среднее значение</i>	<i>Отклонение</i>	<i>Удовлетворяющих проектов, %</i>
<i>CC</i>	2,58	1,13	44
<i>LOCm</i>	15,2	5,72	56
<i>I</i>	0,59	0,11	100
<i>A</i>	0,19	0,15	33

- Не все проекты имеют метрики, удовлетворяющие условию стабильности

# [ Статические метрики: выводы ]

- Статические метрики не информативны
  - Значения специфичны для проекта
  - «Средние» проекты редки
  - Значения сложно интерпретировать

# [ Исторические метрики ]

- Функции времени (ревизии)
- Отображают динамику процесса
- Допускают статистический анализ

# [ Разностные метрики ]

- Разности первого порядка
  - $\Delta_r M = M_{r_2} - M_{r_1}$
  - $\Delta_{(r_1, r_2)} M = (M_{r_2} - M_{r_1}) / (r_2 - r_1)$
- Анализ
  - Знак – направление изменения
  - Величина – масштаб изменения



# Кластеризация

## ■ Кластеризация

- Группировка по направлению, масштабу

	$\Delta CC = 0$	$\Delta CC > 0$	$\Delta CC < 0$
$\Delta LOC = 0$	Косметика	Исправление ошибки	Исправление ошибки, Рефакторинг
$\Delta LOC > 0$	Расширение функции	Новая функциональность	Рефакторинг
$\Delta LOC < 0$	Сужение функции	Исправление ошибки	Удаление функциональности

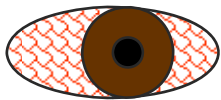
# [ Применение ]

---

- Автоклассификация, фильтрация
- Контроль стадий разработки
- Другое

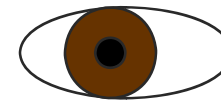
# Классификация изменений

- Концентрация на типах при просмотре:
  - Сокращение времени
  - Улучшение качества



Revision	Actions	Author	Date	Bug ID	Message
23783		tr_builder	22:28:11, 19 октября 2007 г.		Build 4.7.2727
23784		rorov	21:39:45, 19 октября 2007 г.		Исправил сохранение алармов в Alarm
23785		rorov	21:15:52, 19 октября 2007 г.		Чуть изменил модель AlarmInfo Teste;
23785		kluzhev	18:04:34, 19 октября 2007 г.		Поправил реверенсы.
23754		kluzhev	17:52:45, 19 октября 2007 г.	3036	Добавил очень простую (и пока не-
23753		rorov	17:51:09, 19 октября 2007 г.		небольшой редакторчик AlarmModel
23782		rorov	17:46:43, 19 октября 2007 г.		Добавил тесты модели AlarmInfo
23781		rorov	17:44:01, 19 октября 2007 г.		Дизайнил модель AlarmInfo
23780		rorov	17:43:35, 19 октября 2007 г.		Добавил метод UpdateObject в Store
23749		rorov	17:43:13, 19 октября 2007 г.		Добавил комментарий в VersionService
23748		kluzhev	17:42:54, 19 октября 2007 г.		Доработал код новых тестов.
23747		rorov	17:21:40, 19 октября 2007 г.		Изменил схему AlarmInfo
23742		kluzhev	16:57:40, 19 октября 2007 г.		Улучшил реализацию AlarmInfo
23744		kluzhev	16:37:43, 19 октября 2007 г.		Используется еще одна строка для
23741		kluzhev	16:35:16, 19 октября 2007 г.		Вызвали из оболочки тестовой сборки
23739		rorov	15:47:48, 19 октября 2007 г.		Изменил схему AlarmModel
23738		rorov	14:41:06, 19 октября 2007 г.		Изменил схему AlarmModel
23735		rodman	13:18:28, 19 октября 2007 г.		
23731		rodman	10:49:59, 19 октября 2007 г.		
23723		tr_builder	18:55:56, 18 октября 2007 г.		Build 4.7.2726
23722		rorov	18:28:20, 18 октября 2007 г.		Добавил warning NHQR-3043
23720		rorov	18:08:26, 18 октября 2007 г.		Изменил реализацию AlarmSaver'a
23718		rorov	18:06:02, 18 октября 2007 г.		Слегка изменил модель AlarmInfo и т.
23717		rodman	17:56:33, 18 октября 2007 г.		
23716		rorov	17:45:50, 18 октября 2007 г.		Изменил схему
23712		tr_builder	16:26:44, 18 октября 2007 г.		Build 4.7.2725
23711		rorov	15:57:23, 18 октября 2007 г.		Изменил схему
23710		kluzhev	15:52:26, 18 октября 2007 г.		Удале лямбда-исп.
23709		kluzhev	15:31:31, 18 октября 2007 г.		Имя
23707		rorov	15:48:09, 18 октября 2007 г.		Переименовал AlarmConditionSetting
23705		tr_builder	13:58:16, 18 октября 2007 г.		Build 4.7.2724
23704		rorov	13:25:39, 18 октября 2007 г.		Добавил StatusInfoFoundException
23703		rorov	13:13:32, 18 октября 2007 г.		Вынес GVAlarmMessageTranslator в т.
23702		rorov	13:13:00, 18 октября 2007 г.		Вынес GVMessageQueueBase в отде
23701		rorov	13:12:31, 18 октября 2007 г.		Переименован GVMessageTranslator
23700		rorov	13:11:17, 18 октября 2007 г.		Переименован GVMessageLoader в c
22699		tr_builder	13:09:40, 18 октября 2007 г.		Build 4.7.2723

```
1 namespace-Monitoring-Model http://svn.traxxas.com/savimas/NBQR4/trunk/
2
3 .....public-abstract-class-AlarmCondition
4 .....public-abstract-class-AlarmCond
```



```
1 namespace-Monitoring-Model http://svn.traxxas.com/savimas/NBQR4/trunk/
2
3 .....public-abstract-class-AlarmCondition
4 .....public-abstract-class-AlarmCond
5 .....private-string-message
6 .....private-nullable-date-time-
7 .....private-int-message-type
8 .....public-string-message
9 .....private-nullable-date-time-
10 .....private-int-message-type
11 .....private-int-message-type
12 .....public-string-message
13 .....public-string-message
14 .....public-string-message
15 .....public-string-message
16 .....public-string-message
17 .....public-string-message
18 .....public-string-message
19 .....public-string-message
20 .....public-string-message
21 .....public-string-message
22 .....public-string-message
23 .....public-string-message
24 .....public-string-message
25 .....public-string-message
26 .....public-string-message
27 .....public-string-message
28 .....public-string-message
29 .....public-string-message
30 .....public-string-message
31 .....public-string-message
32 .....public-string-message
33 .....public-string-message
34 .....public-string-message
35 .....public-string-message
36 .....public-string-message
37 .....public-string-message
38 .....public-string-message
39 .....public-string-message
40 .....public-string-message
41 .....public-string-message
42 .....public-string-message
43 .....public-string-message
44 .....public-string-message
45 .....public-string-message
46 .....public-string-message
47 .....public-string-message
48 .....public-string-message
49 .....public-string-message
50 .....public-string-message
51 .....public-string-message
52 .....public-string-message
53 .....public-string-message
54 .....public-string-message
55 .....public-string-message
56 .....public-string-message
57 .....public-string-message
58 .....public-string-message
59 .....public-string-message
60 .....public-string-message
61 .....public-string-message
62 .....public-string-message
63 .....public-string-message
64 .....public-string-message
65 .....public-string-message
66 .....public-string-message
67 .....public-string-message
68 .....public-string-message
69 .....public-string-message
70 .....public-string-message
71 .....public-string-message
72 .....public-string-message
73 .....public-string-message
74 .....public-string-message
75 .....public-string-message
76 .....public-string-message
77 .....public-string-message
78 .....public-string-message
79 .....public-string-message
80 .....public-string-message
81 .....public-string-message
82 .....public-string-message
83 .....public-string-message
84 .....public-string-message
85 .....public-string-message
86 .....public-string-message
87 .....public-string-message
88 .....public-string-message
89 .....public-string-message
90 .....public-string-message
91 .....public-string-message
92 .....public-string-message
93 .....public-string-message
94 .....public-string-message
95 .....public-string-message
96 .....public-string-message
97 .....public-string-message
98 .....public-string-message
99 .....public-string-message
100 .....public-string-message
```

# Автоклассификация изменений

- **Входные данные**
  - Классы изменений
  - Набор метрик
- **Инициализация на обучающем множестве**
  - Расчет метрик
  - Кластеризация
  - Экспертная интерпретация
- **Автоклассификация основного множества**
- **Выходные данные**
  - Изменения распределены по классам

# [ Классы изменений ]

---

- Новая функциональность
- Рефакторинг
- Удаление кода
- Исправление ошибки
- Косметическое изменение
- Другие

# Метрики изменений

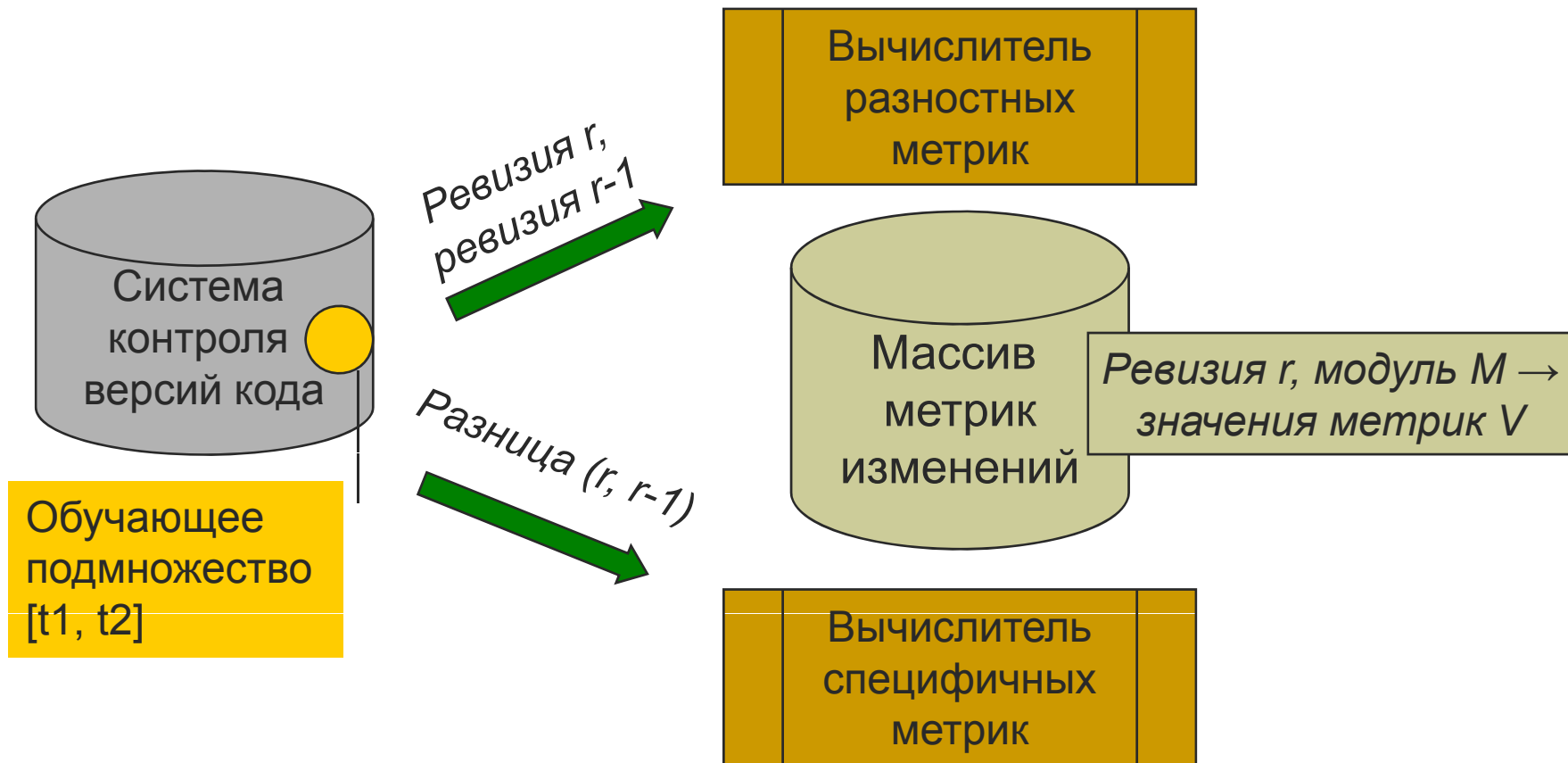
- Расширение стандартных метрик:

- $\Delta_r CC = CC_{r2} - CC_{r1}$
- $\Delta_r C/S = C/S_{r2} - C/S_{r1}$
- $\Delta_r eLOC = eLOC_{r2} - eLOC_{r1}$
- ...

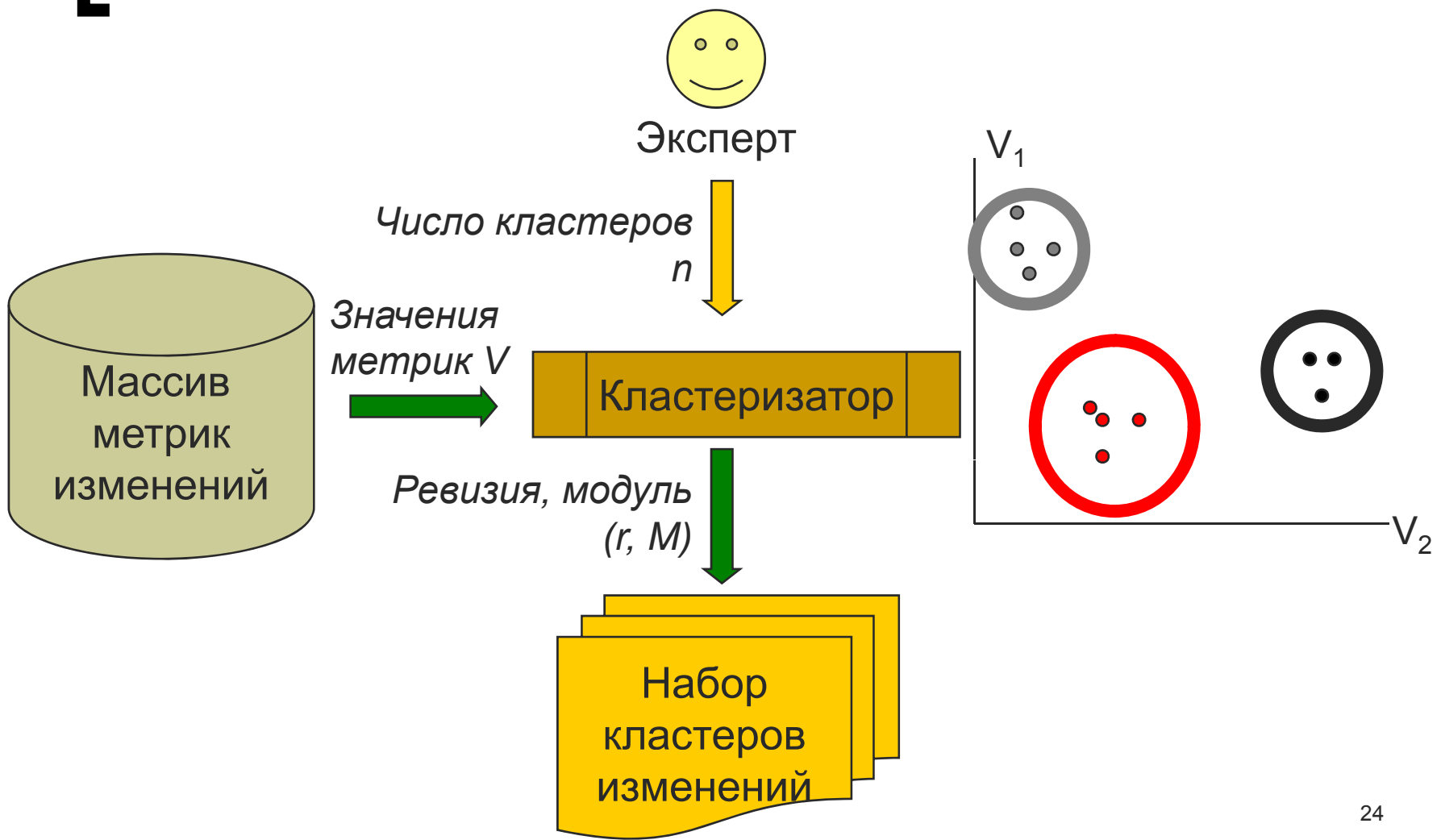
- Специфичные метрики:

- $CH\_LOC(r) = eLOC(Diff(r))$
- ...

# Расчет метрик обучающего множества

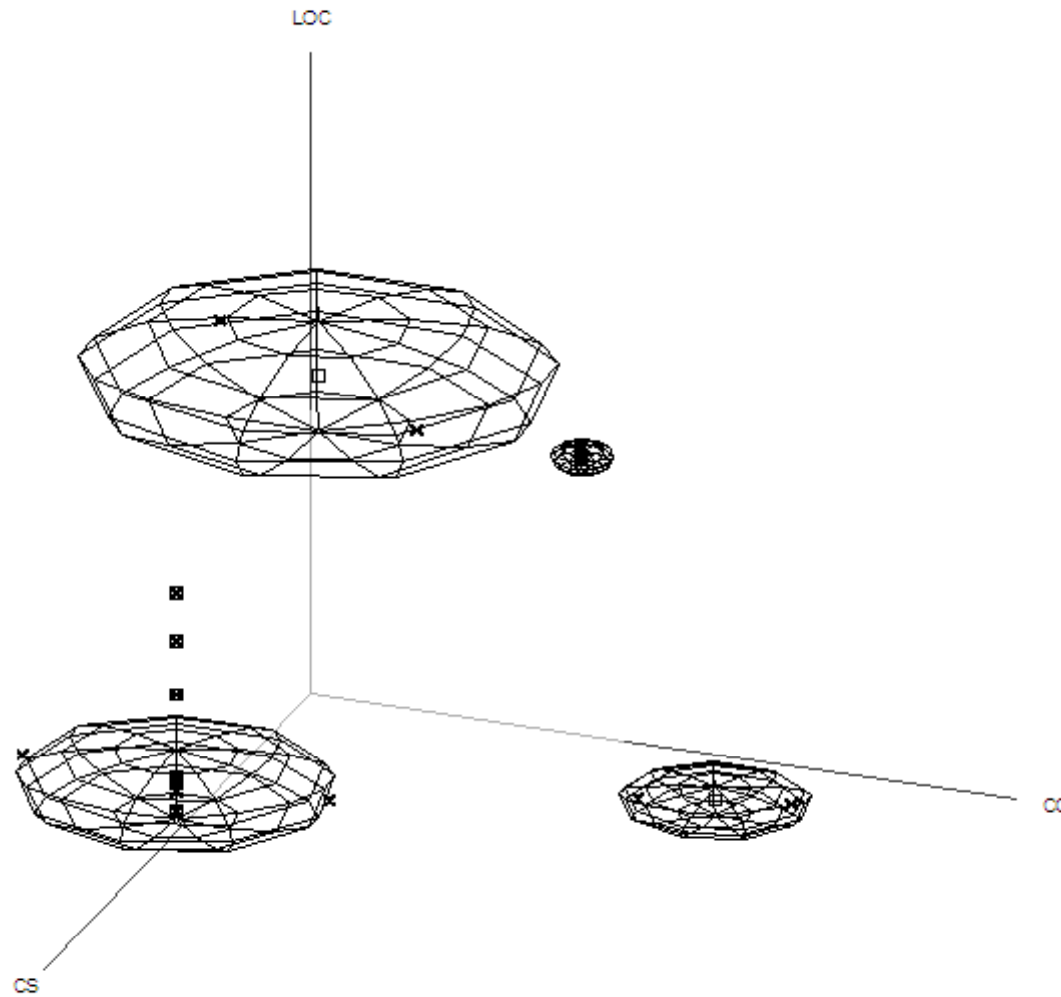


# Кластеризация обучающего множества

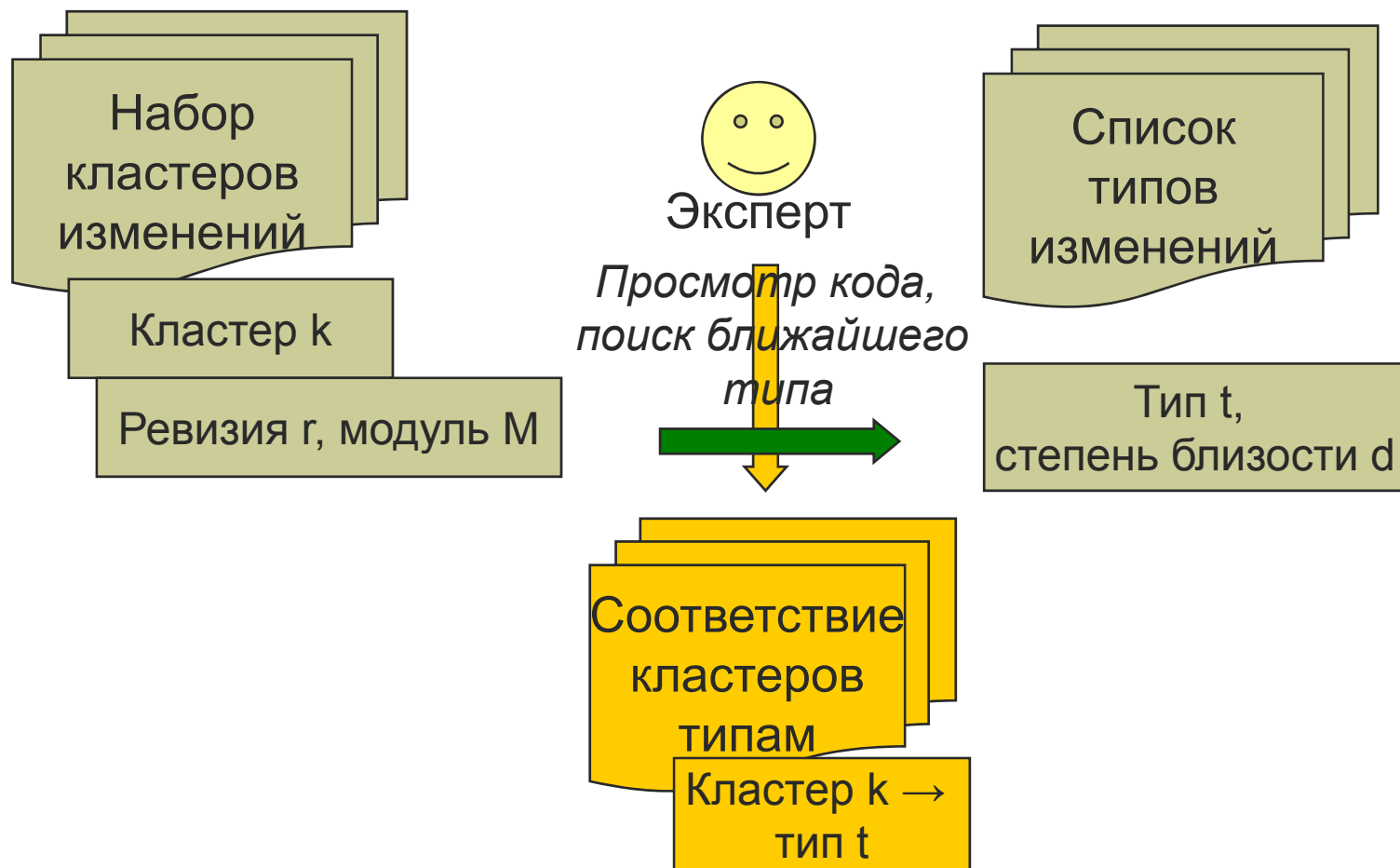




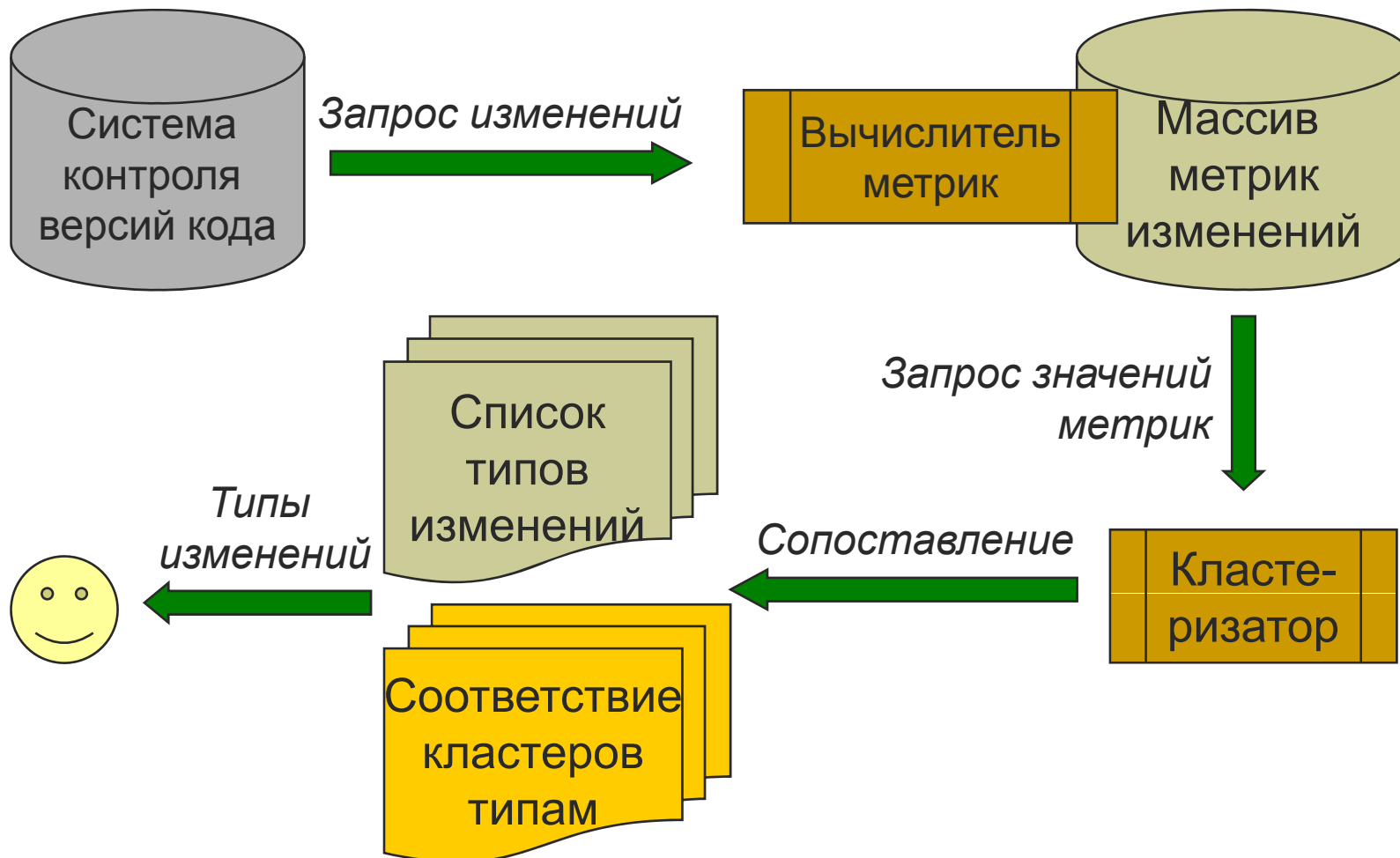
# [ Результат кластеризации: 3D ]



# Экспертная интерпретация

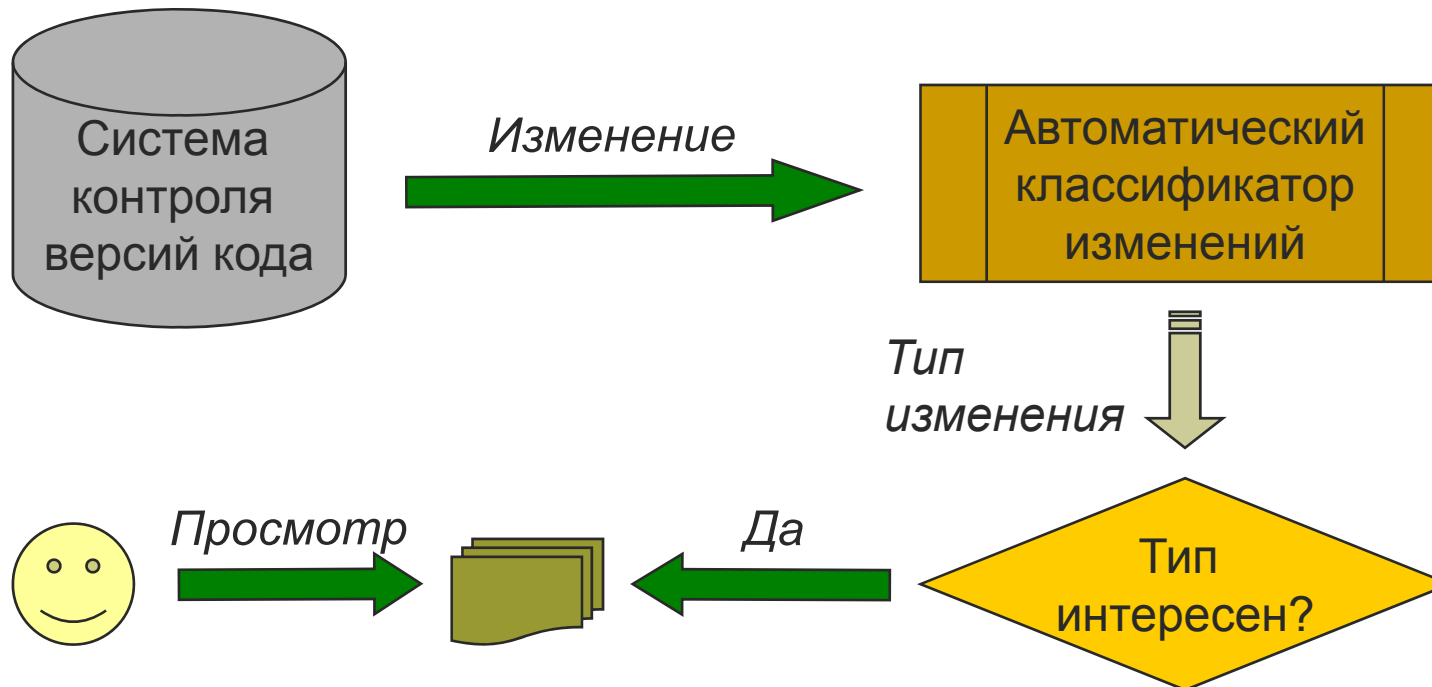


# Классификация основного множества



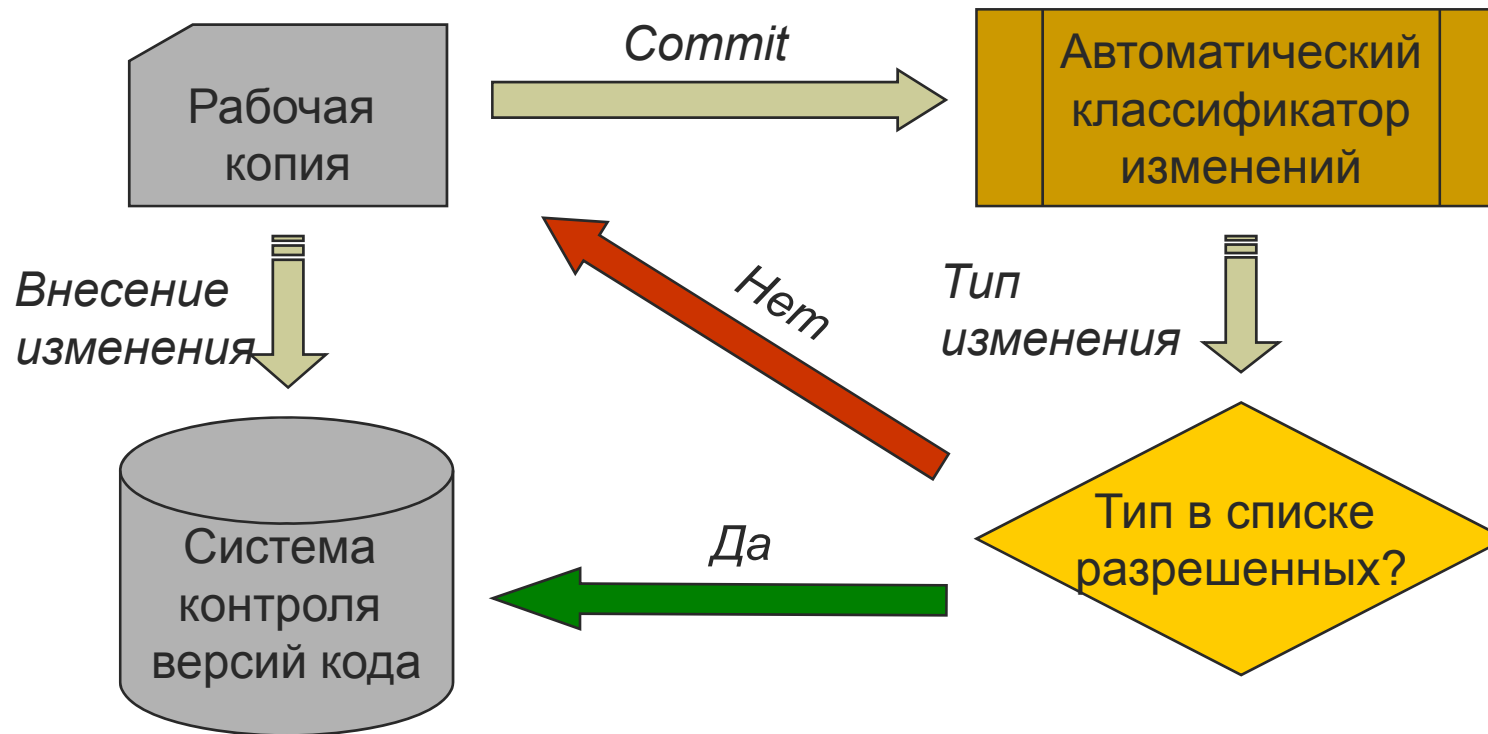
# [ Разработчик ]

- Просмотр изменений заданного типа



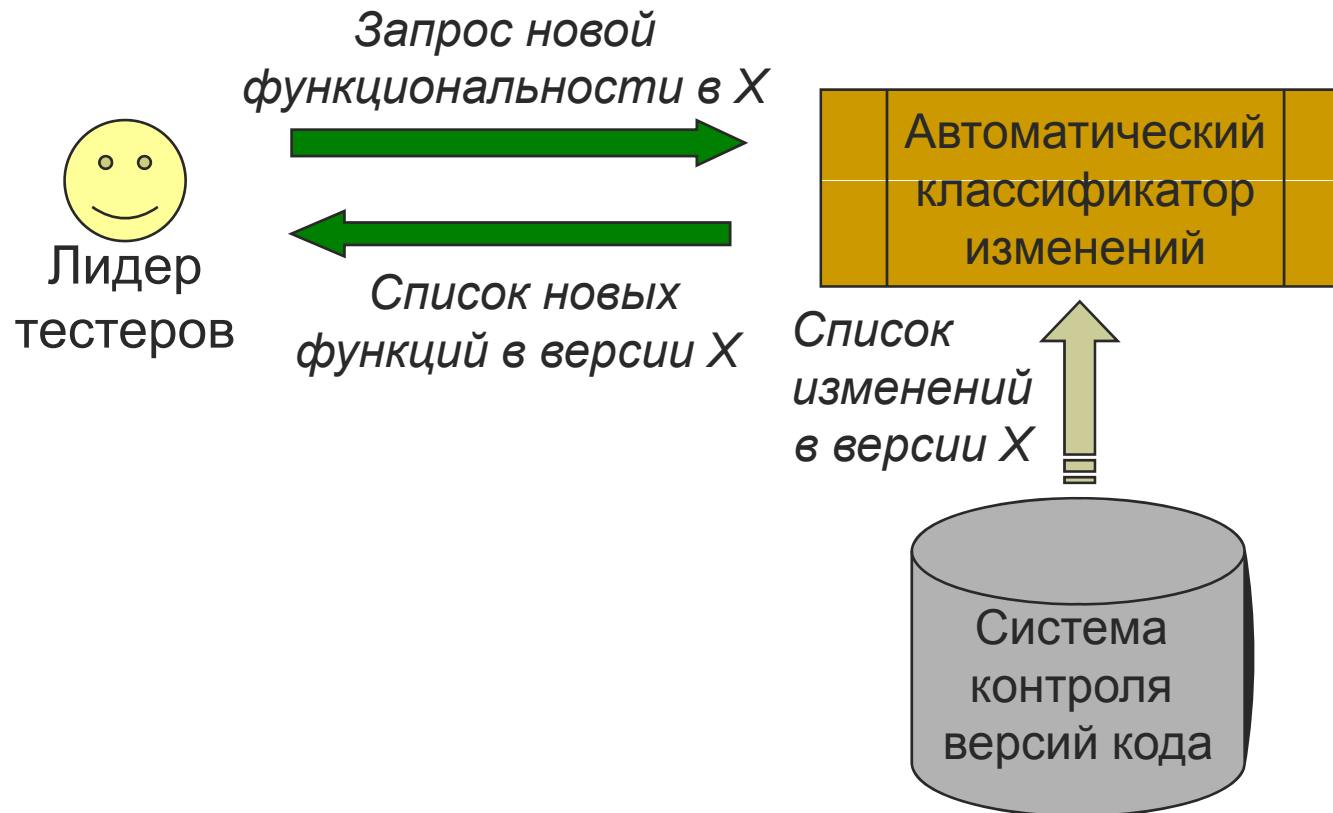
# [ Лидер разработчиков ]

- Запрет некоторых типов изменений



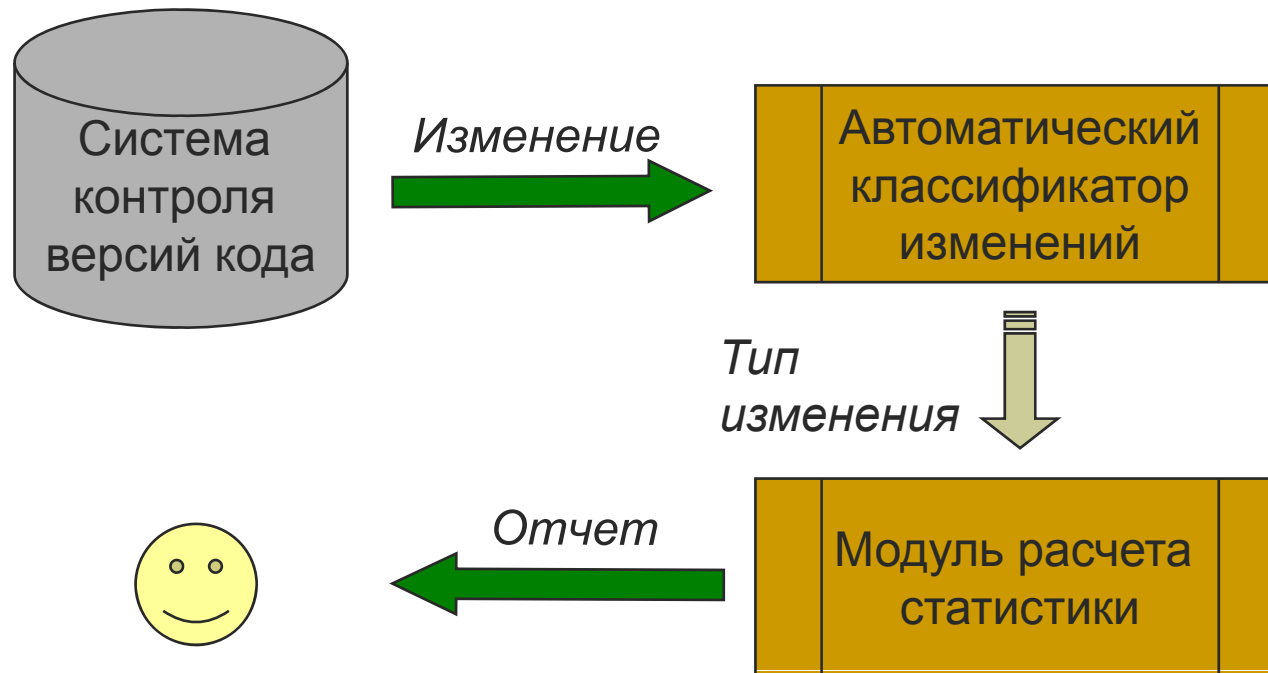
# Лидер тестеров

- Получение списка функциональности для конкретной версии



# [ Менеджер проекта ]

- Анализ статистики изменений



# Статистика изменений

<i>Проект</i>	<i>Navi-Manager</i>	<i>TortoiseSVN</i>
<i>Класс изменения</i>		
<i>Все</i>	72	215
<i>Новая функциональность</i>	4	73
<i>Рефакторинг</i>	16	60
<i>Удаление</i>	1	0
<i>Исправл. ошибок</i>	50	82
<i>Другие</i>	1	0



# Результат классификации

Ревизия	СС	IC	eLOC	Эксперт. классиф.	Авт. класс.	Комментарий к изменению
16820	-2	+1	-5	рефакт.	1	Процесс объединения кораблей в рамках одной транзакции.
16833	0	0	-1	удал.фун.	2	Удалена лишняя команда (коммит).
17026	+4	+4	+18	нов.фун.	3	Вывод широты, долготы в полном формате при поллинге.
17029	0	0	+1	правка ошибки	4	Обновление MessageSource. UpdateTime после обработки каждого сообщения.
17038	0	0	-1	удал.фун.	2	Откат изменений из r17029. Не нужно обновлять поле после обработки каждого сообщения.
17107	+4	+12	+89	нов.фун.	3	Первая версия механизма загрузки треков с MonServer.

# [ Достоинства метода ]

---

- Универсальность
- Формальность
- Настраиваемость

# [ Ограничения метода ]

- Смешанные типы изменений
- Новые типы в процессе классификации
- Для каждого проекта нужно обучать

# [ Будущие исследования ]

- Анализ составных типов изменений
- Нечеткий метод кластеризации (fuzzy k-Means)
- Объектные и другие метрики
- Оценивание применимости для различных типов проектов

# [ Вопросы ]

---

[ Спасибо за внимание! ]

---

- [evgeny.knyazev@gmail.com](mailto:evgeny.knyazev@gmail.com)
- [danil.shopyrin@gmail.com](mailto:danil.shopyrin@gmail.com)