

СИСТЕМАТИЧЕСКИЙ ПОДХОД К РАЗРАБОТКЕ ТЕСТОВЫХ ЗАДАНИЙ (Evidence Centered Design – ECD)

*Концептуальные подходы к разработке и
скоринг инструмента оценки ИК-
компетентности*

Руднев Максим
кандидат соц.наук, эксперт
Национального фонда подготовки
кадров

Тарасова Ксения
кандидат пед.наук, эксперт
Национального фонда подготовки
кадров



Роберт Мислеви (Robert Mislevy)

Разработчик
систематического подхода к
разработке тестовых
заданий (Evidence Centered
Design – ECD)

В последнее время -
Principled Assessment Design
for Inquiry in science – PADI

Представление об измерении латентных переменных (о предмете психометрике)

1) Прагматическое (классическое) модернистское – от Платона: идеи существуют, но не наблюдаемы, тени – лишь тени, зато наблюдаемы и по ним можно пытаться угадать идеи – true score.

Развитие: не все тени одинаково полезны (DIF, IRT и т.д.)

2) Постмодернистское: стандартизованные тесты не подходят всем одинаково. Следовательно, балл не имеет общего для всех смысла.

Развитие: адаптивное тестирование, формативное оценивание. Доверие?

3) Неопрагматическое постмодернистское: стандартизованные тесты имеют смысл, но только в рамках определенного дискурса. Осознание границ дискурса делает применимым достижения прагматического подхода.

Дискурс складывается из понимания самого латентного признака, его связи с другими концептами в социальном контексте, а также ситуации и способов измерения.

Критерий валидности – не степень соответствия «истине», а соответствие условным конвенциям.

Систематический подход к разработке теста (ECD) – это набор взаимосвязанных процедур, призванных прояснить два вопроса:

- 1. Какие свидетельства компетентности (или знаний) учащегося мы можем непосредственно наблюдать?*
- 2. Как структурировать ситуацию, чтобы увидеть эти свидетельства?*

ECD – это язык, на котором описывается процесс разработки теста

Классический	ECD
Уровень подготовленности/знаний/компетентности	Гипотетическое утверждение (Claim)
--	Наблюдаемые переменные
Ответы тестируемого	Результаты деятельности в структурированной ситуации
Формат вопроса	Контент, контекст, ситуация, средства презентации, средства администрирования, допущения о знаниях/компетенциях из других областей
Результаты тестирования	Аргументы в пользу утверждения

СИСТЕМАТИЧЕСКИЙ ПОДХОД К РАЗРАБОТКЕ ТЕСТА

1. Анализ области компетенций



2. Моделирование области компетенций



3. Теоретическая рамка теста



4. Воплощение теста



5. Применение теста

Что мы знаем об этой области? В каких ситуациях она важна? Как они проявляются? Каковы ключевые компетенции? (иногда называли Job analysis)

Как мы концептуализируем ключевые аспекты области компетенций? Сколько суб-компетенций? Как они связаны между собой?

Каковы основные элементы теста? Модель студента, свидетельства и задания.

Наполнение основных элементов теста содержанием: создание конкретных заданий, статистические модели

Взаимодействие участников тестирования с тестом, оценка компетенций, обратная связь.

1. Анализ области компетенций (Domain analysis)

Что мы знаем об этой области?

*Какой подход к рассмотрению выбираем
(когнитивный, бихевиористский,
интеракционистский и т.п.)*

В каких ситуациях она важна?

Как они проявляются?

Каковы ключевые компетенции?

Job analysis

User experience

(Внешние) причины и следствия компетенций

2. Моделирование области компетенций (Domain modeling)

Как мы концептуализируем ключевые аспекты области компетенций?

Сколько суб-компетенций?

Как они связаны между собой?

Как связаны элементы ситуаций, в которых компетенция может проявиться?

1. Определения компетенций (student model)

Что мы измеряем?

Определение объема измеряемой компетентности и ее составляющих, каждого из уровней либо смысла баллов и способов описания каждого из уровней

2. Модель сбора свидетельств (evidence model)

Как мы можем это измерить?

Поиск наблюдаемых свидетельств наличия компетентности

3. Модель тестового задания (task model)

Где измерять?

Поиск конкретных индикаторов – результатов выполнения какой-либо задачи.

3. Теоретическая рамка теста

4. Правила набора заданий (assembly model)

Сколько и каких заданий требуется?

Критерии отбора заданий для конкретного тестирования.

5. Выбор средств презентации заданий (presentation model)

В каком виде представлять задания?

Решение о формате теста: бумага/компьютер, множественный выбор/виньетки/сценарий.

6. Способы администрирования (delivery model)

Какие средства и технологии использовать для представления заданий и сбора ответов?

Варианты обеспечения администрирования: как доставить задание, как записать ответ, как обработать результат (правила скоринга), вернуть ли результат тестируемому, как записать результат и мета-информацию для анализа заданий.

*=> Набор конкретных ответов по этим шести пунктам дает
Спецификацию теста*

4. Воплощение теста

Реализация спецификации теста.

Например:

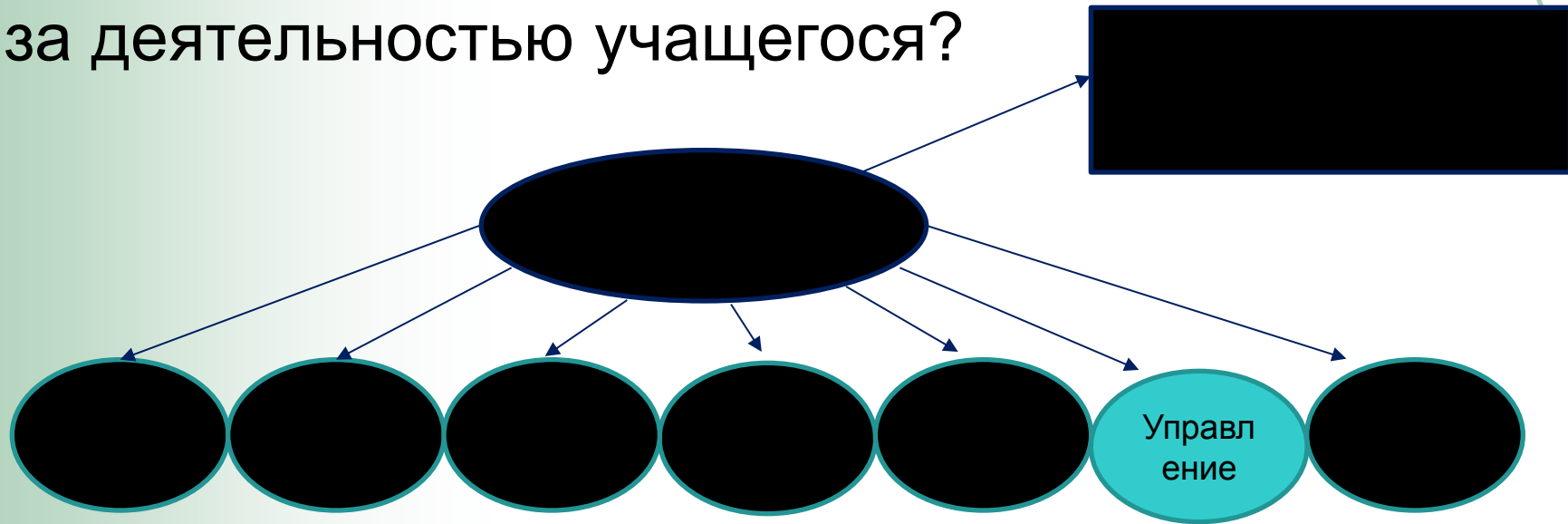
1. **Определения компетенций (student model)** – многомерный дискретный конструкт ИКК, отдельная компетенция – Управление. Управление на высоком уровне – (в числе прочего) успешно создает или выбирает схему классификации.
2. **Модель сбора свидетельств (evidence model)** – выбор наблюдаемых переменных, конкретных индикаторов: управление рассылками
3. **Модель тестового задания (task model) - *контент***, т.е. конкретное содержание тестовых заданий (например, проблемная ситуация засоренного спамом почтового ящика). И ***контекст***, т.е. тип ситуации решения проблемы (например, учебный или бытовой)
4. **Правила набора заданий (assembly model)** – количество и список характеристик тестовых заданий, которые должны присутствовать в каждом варианте теста.
5. **Выбор средств презентации заданий (presentation model)** – сценарное задание на экране компьютера, симулирующее веб-интерфейс электронной почты
6. **Способы администрирования (delivery model)** – локально-сетевой способ администрирования, в присутствии администратора, самостоятельное, результаты деятельности, включая контекстуальные, автоматически записываются на сервер через интернет-соединение.

■ 1. Модель компетенций: что измеряем и зачем?

ИКК = многомерный дискретный конструкт



- Модель сбора свидетельств: что мы можем узнать о компетенции управления, наблюдая за деятельностью учащегося?

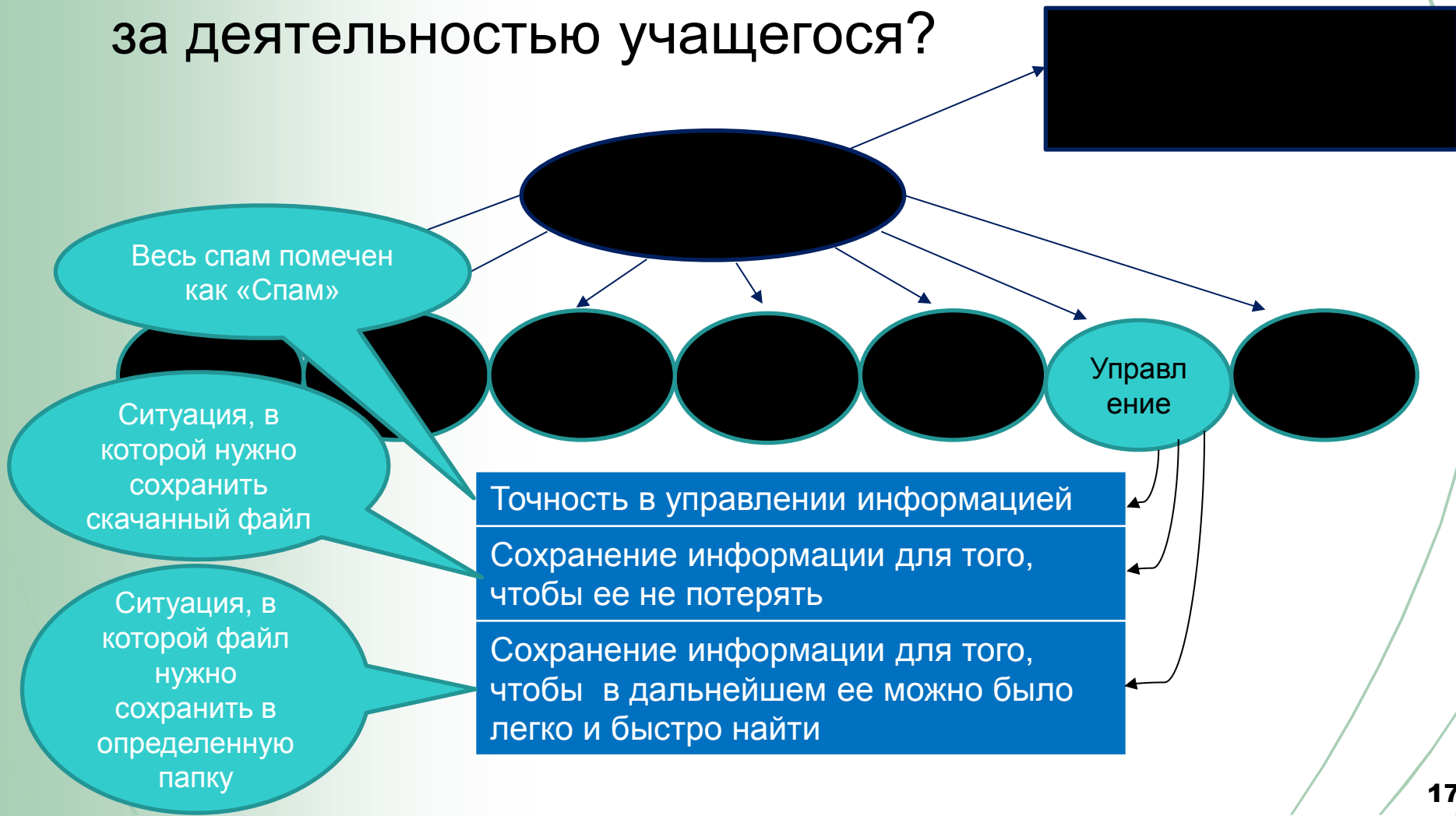


Управление на высоком уровне – (в числе прочего) успешно создает или выбирает схему классификации.

- Модель сбора свидетельств: что мы можем узнать о компетенции управления, наблюдая за деятельностью учащегося?



- Модель сбора свидетельств: что мы можем узнать о компетенции управления, наблюдая за деятельностью учащегося?



- Модель тестового задания: как мы можем структурировать ситуацию, чтобы увидеть свидетельства компетентности?

Тестовое задание

+контекст

+контент

Весь спам помечен
как «Спам»

Ситуация, в
которой нужно
сохранить
скачанный файл

Ситуация, в
которой файл
нужно
сохранить в
определенную
папку

Точность в управлении информацией

Сохранение информации для того,
чтобы ее не потерять

Сохранение информации для того,
чтобы в дальнейшем ее можно было
легко и быстро найти

Управл
ение

- Одно из важнейших преимуществ теста, построенного на системе свидетельств – обратимость логики. Эти процедуры «автоматически» повышают конструктивную валидность теста.
- *Пример: в ситуации необходимости сортировки спама, в которой учащийся может проявить свою компетентность таким образом, чтобы мы ее смогли зафиксировать, учащийся отсортировал спам, что является одним из свидетельств его компетентности в части управления информацией, и, следовательно, одним из проявлений общей ИК-компетенции.*

- Связи между элементами компетенций, условиями сбора свидетельств и моделями тестовых заданий может быть крайне сложной.
- Байесовские сети позволяют моделировать любой уровень сложности структуры, поэтому они особенно хорошо подходят для обсчета результатов теста, основанного на ECD. (Об этом позже)

5. Применение теста

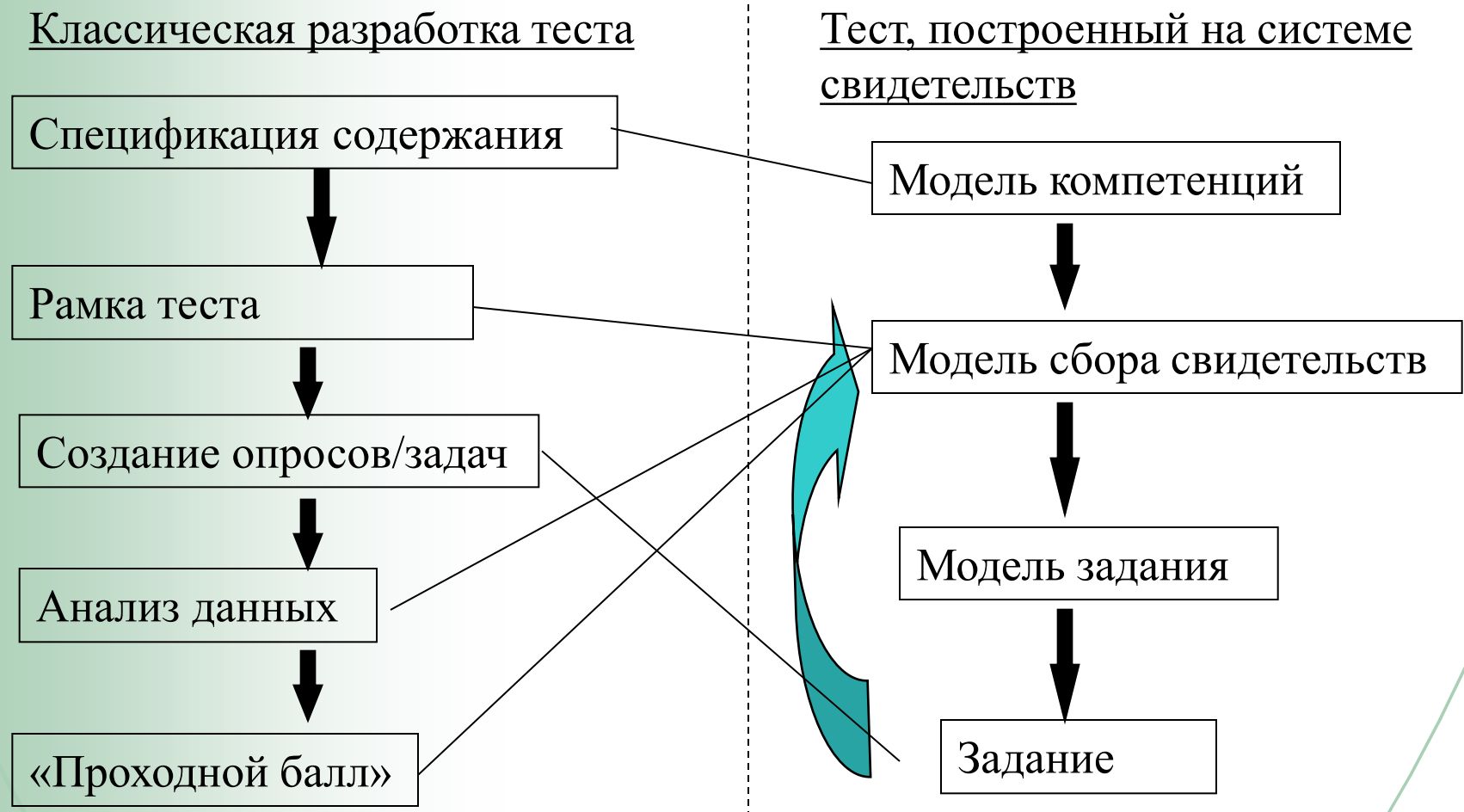
Взаимодействие участников тестирования, администраторов и учителей с тестом.

Сбор данных.

Оценка компетенций (включая статистическую модель).

Обратная связь (с кем?).

Процедуры разработки теста в двух разных подходах



- Для более подробного знакомства с ECD
несколько статей можно скачать
<http://smarturl.co/ECD>

Пример применения подхода ECD в рамках ICL Test

Общие параметры задания
«Сортировка электронной почты»

Предмет оценки:
«Управление (информацией)»
(успешно создает или выбирает схему классификации, организации)
Продолжительность: 4 минуты

Контекст: личный

Содержание: практические задачи

Задача: сохранить информацию из писем таким образом, чтобы ее не потерять, чтобы в дальнейшем ее можно было найти легко и быстро.

Ты только что вернулся домой после летних каникул. Решив проверить свою почту, ты обнаружил в папке «Входящие» сообщение о том, что почтовый ящик переполнен и что ты не сможешь отправлять и получать новую почту до тех пор, пока не увеличишь свободное место в папке «Входящие». Ты знаешь, что можно спокойно удалить рекламные письма с пометкой Спам, а также все письма до 15 ноября, касающиеся Нельсоновского проекта. У тебя нет времени, чтобы прочитать все письма и вложения в почтовом ящике, поэтому ты решил сначала сохранить пять самых больших вложений в соответствующие папки на жестком диске ('D:\Вложения'), а затем стереть пять писем с этими вложениями.

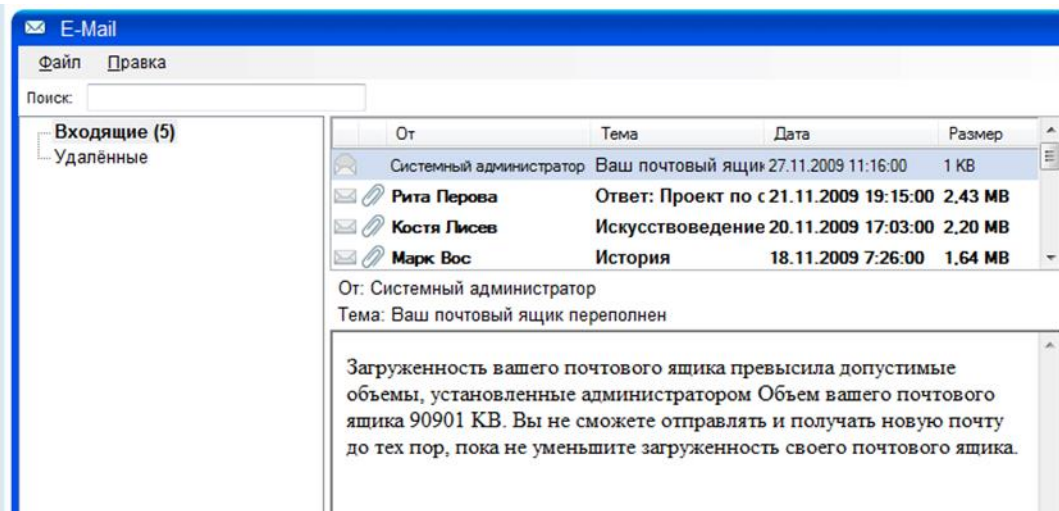
Щелкни по кнопке ДАЛЕЕ, чтобы начать выполнение задания.

Инструкции

Чтобы освободить место в папке «Входящие» тебе нужно:

- Удалить рекламные письма с пометкой «Спам»
- Удалить письма, полученные до 15 ноября, касающиеся Нельсоновского проекта, в поле «Тема»
- Сохранить пять самых больших вложений в соответствующие папки на жестком диске ('D:\Вложения')
- Удалить пять писем с самыми большими вложениями, которые ты только что сохранил

Щелкни по кнопке ДАЛЕЕ, когда закончишь выполнение задания.



Модель сбора свидетельств

Как мы можем это измерить: поиск наблюдаемых свидетельств наличия компетентности (что мы можем узнать о компетенции управления, наблюдая за деятельностью испытуемого)

Наблюдаемые признаки/Измеряемые переменные (Observables)	Уровень
Точность в управлении информацией	Высокий: удалено 10 ненужных писем (4 письма с пометкой «СПАМ» и 1 письмо с пометкой Nelson, 5 больших по объему писем) и не удалены (оставлены) 3 нужных письма. Средний: Работа выполнена на 80%/Обработано 80% писем Низкий: Работа выполнена менее чем на 80%
Сохранение информации для того, чтобы ее не потерять	Высокий: сохранены все электронные письма, которые необходимо сохранить. Средний: сохранено 80% электронных писем, которые необходимо сохранить. Низкий: сохранено менее 80% электронных писем, которые необходимо сохранить.
Сохранение информации для того, чтобы в дальнейшем ее можно было легко и быстро найти	Высокий: сохранены 5 файлов в соответствующие папки Средний: сохранены 4 файла в соответствующие папки Низкий: сохранено менее 4 файлов в соответствующие папки

Этап 1. Что измеряем

- выбор одной или нескольких компетенций
- операционализация определения компетенции/компетенций

Этап 2. Как измеряем

В соответствии со спецификацией теста выбираем:

- продолжительность выполнения тестового задания (4, 15, 30 минут)
- содержание (реалистичная ситуация, заключающая в себе проблему, интересная для решения: гуманитарные науки, общественные науки, естественные науки, поп-культура, практические задачи)
- контекст (учебный или личный)
- основные инструменты (программные средства)

Этап 3. Выделение измеряемых переменных

Этап 4. Написание сценария тестового задания

Спецификация теста

Содержание	Контекст	Основные инструменты
HUMANITIES (2)	ACADEMIC (8)	Email (6):
SOCIAL STUDIES (3)		Word Processing (2+1)
NATURAL SCIENCES (3)		Presentation (1+1):
PRACTICAL AFFAIRS (5)	PERSONAL (8)	Search Strategies (2)
POPULAR CULTURE (3)		Spreadsheet (+1)
		Graph (1)
		Community (1+1)
		IM (3)
		Web page (3)
		Table (1)
		Concept Map (1)
		Other (1)

Уровень сложности задания	Количество заданий	Ожидаемое время выполнения задания
Простое (оценка 1-й составляющей ИК-компетентности)	13	3 - 4
Средней сложности (оценка 2-3 составляющих ИК-компетентности)	2	10 -15
Сложное (оценка 4-5 составляющих ИК-компетентности)	1	20 - 30

Широкий спектр цифровых технологий: текстовые редакторы, инструменты для создания презентаций, электронные таблицы, графики, базы данных, средства мультимедиа, электронная почта, социальные сети и др. Интернет-сервисы.

Пример разработки тестового задания: «Сортировка электронной почты»

1. Что измеряем

- *компетенция:* Управление (определение: применение существующей организационной или классификационной схемы)
- *операционализация:* предпринять элементарную и предварительную организацию имеющейся информации для дальнейшего использования

2. Как измеряем

В соответствии со спецификацией теста выбираем:

- продолжительность выполнения: 4 минуты
- содержание: практические задачи
- контекст: личный
- основные инструменты: электронная почта

Пример разработки тестового задания: «Сортировка электронной почты»

3. Определение измеряемых переменных

- Удаление писем в соответствии с заданным критерием (точность в управлении информацией)
- Сохранение вложений в соответствии с заданным критерием (сохранение информации для того, чтобы ее не потерять)
- Распределение файлов вложений по правильным папкам (сохранение информации для того, чтобы в дальнейшем ее легко и быстро можно было найти)

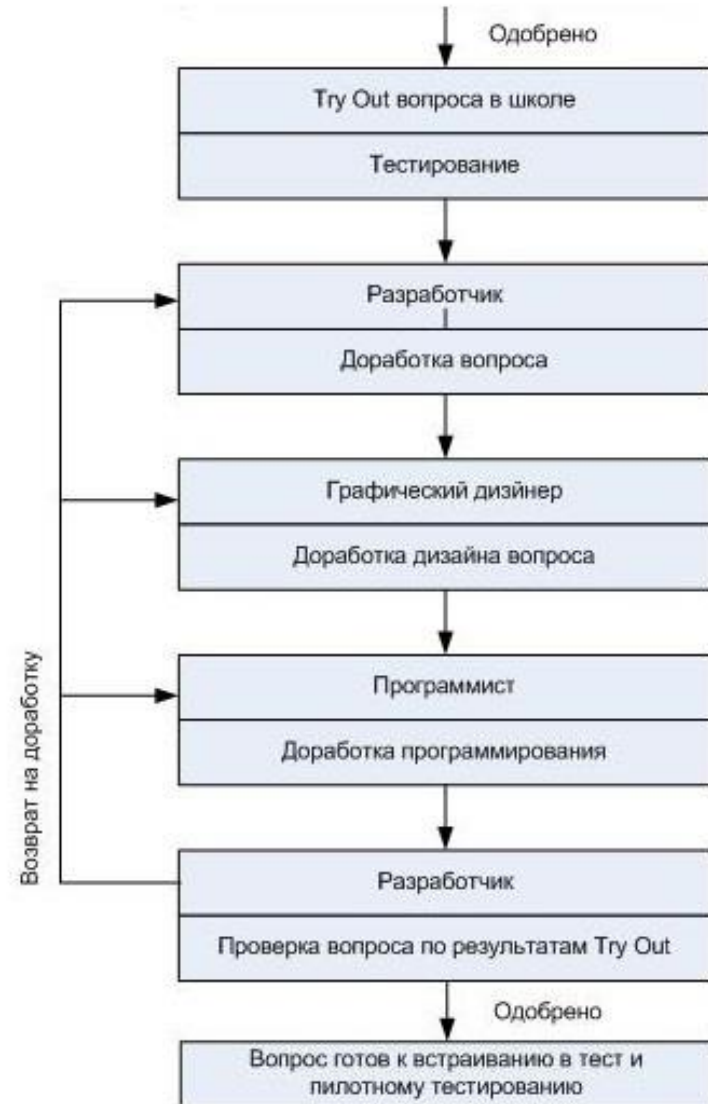
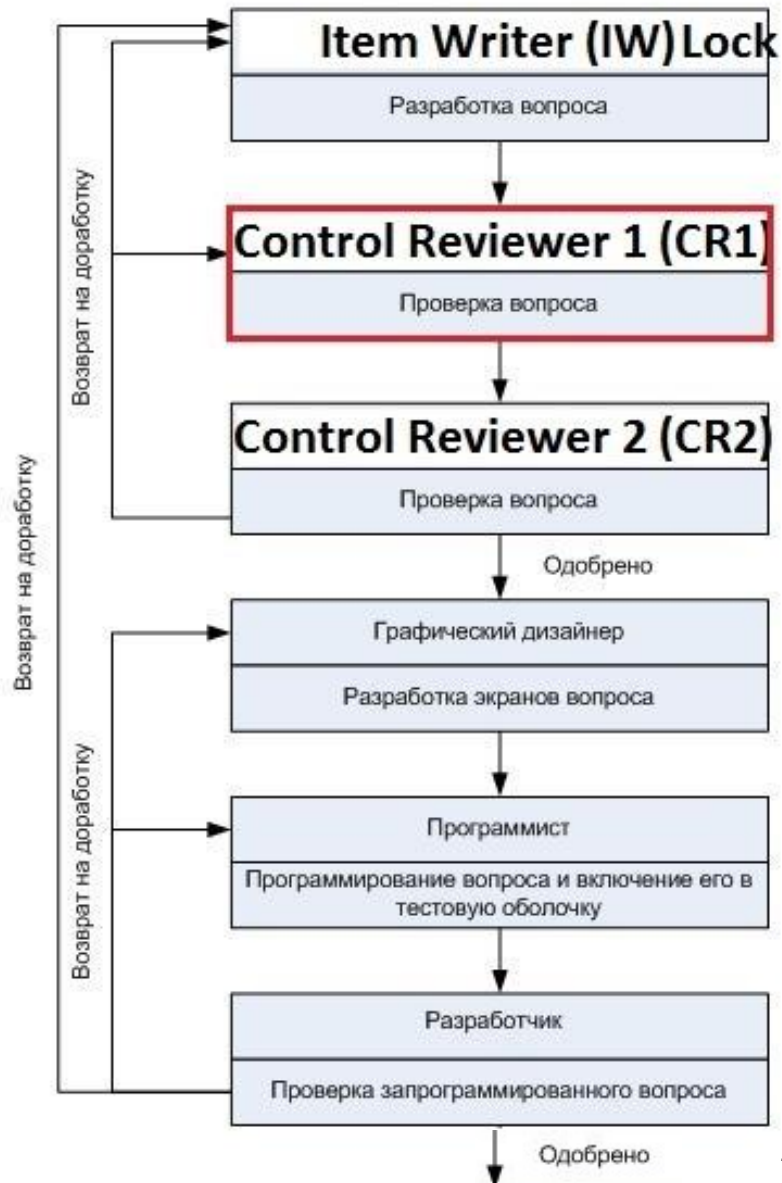
4. Написание сценария задания

- Проблема: ваш почтовый ящик переполнен.
Сохраните самые большие вложения в соответствующие папки, удалите самые большие и ненужные (с пометкой «СПАМ» или «Нельсон» в теме письма) письма.

5. Разработка дизайна

6. Программирование

Пример разработки тестового задания: организационная схема



Ты только что вернулся домой после летних каникул. Решив проверить свою почту, ты обнаружил в папке «Входящие» сообщение о том, что почтовый ящик переполнен и что ты не сможешь отправлять и получать новую почту до тех пор, пока не увеличишь свободное место в папке «Входящие». Ты знаешь, что можно спокойно удалить рекламные письма с пометкой Спам, а также все письма до 15 ноября, касающиеся Нельсоновского проекта. У тебя нет времени, чтобы прочитать все письма и вложения в почтовом ящике, поэтому ты решил сначала сохранить пять самых больших вложений в соответствующие папки на жестком диске ('D:\Вложения'), а затем стереть пять писем с этими вложениями.

Щелкни по кнопке ДАЛЕЕ, чтобы начать выполнение задания.

Инструкции

Чтобы освободить место в папке «Входящие» тебе нужно:

- Удалить рекламные письма с пометкой «Спам»
- Удалить письма, полученные до 15 ноября, касающиеся Нельсоновского проекта, в поле «Тема»
- Сохранить пять самых больших вложений в соответствующие папки на жестком диске ("D:\Вложения")
- Удалить пять писем с самыми большими вложениями, которые ты только что сохранил

Щелкни по кнопке ДАЛЕЕ, когда закончишь выполнение задания.

✉ E-Mail

Файл Правка

Поиск:

Входящие (5)

Удалённые

	От	Тема	Дата	Размер
	Системный администратор	Ваш почтовый ящик	27.11.2009 11:16:00	1 KB
	Рита Перова	Ответ: Проект по с	21.11.2009 19:15:00	2,43 MB
	Костя Лисев	Искусствоведение	20.11.2009 17:03:00	2,20 MB
	Марк Вос	История	18.11.2009 7:26:00	1,64 MB

От: Системный администратор

Тема: Ваш почтовый ящик переполнен

Загруженность вашего почтового ящика превысила допустимые объемы, установленные администратором. Объем вашего почтового ящика 90901 KB. Вы не сможете отправлять и получать новую почту до тех пор, пока не уменьшите загруженность своего почтового ящика.

Время: 01:57:33

Задание 1 из 1



Далее

Инструкции

Чтобы освободить место в папке «Входящие» тебе нужно:

- Удалить рекламные письма с пометкой «Спам»
- Удалить письма, полученные до 15 ноября, касающиеся Нельсоновского проекта, в поле «Тема»
- Сохранить пять самых больших вложений в соответствующие папки на жестком диске ('D:\Вложения')
- Удалить пять писем с самыми большими вложениями, которые ты только что сохранил

Щелкни по кнопке ДАЛЕЕ, когда закончишь выполнение задания.

✉ E-Mail

Файл Правка

Поиск:

Входящие (4)

Удалённые

	От	Тема	Дата	Размер
	Системный администратор	Ваш почтовый ящик	27.11.2009 11:16:00	1 KB
	Рита Перова	Ответ: Проект по с	21.11.2009 19:15:00	2,43 MB
	Костя Лисев	Искусствоведение	20.11.2009 17:03:00	2,20 MB
	Марк Вос	История	18.11.2009 7:26:00	1,64 MB

От: Костя Лисев

Тема: Искусствоведение

Вложенных файлов: 1, размер: 2,20 MB



Сохранить вложения

В приложении 2 картинки, которые я предлагаю использовать. Перешли их Семену, пожалуйста. У меня нет его адреса. Заранее спасибо Увидимся в пятницу

Время: 01:24:01

Задание 1 из 1

?

Далее

Инструкции

Чтобы освободить место в папке «Входящие» тебе нужно:

- Удалить рекламные письма с пометкой «Спам»
- Удалить письма, полученные до 15 ноября, касающиеся Нельсоновского проекта, в поле «Тема»
- Сохранить пять самых больших вложений в соответствующие папки на жестком диске ("D:\Вложения")
- Удалить пять писем с самыми большими вложениями, которые ты только что сохранил

Щелкни по кнопке ДАЛЕЕ, когда закончишь выполнение задания.

✉ E-Mail

Файл Правка

Поиск:

Входящие (4)

Удалённые

	От	Тема	Дата	Размер
✉	Системный администратор	Ваш почтовый ящик	27.11.2009 11:16:00	1 KB
✉	Рита Перова	Ответ: Проект по с	21.11.2009 19:15:00	2,43 MB
✉	Костя Лисев	Искусствоведение	20.11.2009 17:03:00	2,20 MB
✉	Марк Вос	История	18.11.2009 7:26:00	1,64 MB

Сохранить файл

D:\Вложения

- Искусствоведение
- История
- Нельсон
- Социология
- Фото

Имя файла: Искусство.jpeg

Сохранить

Отмена

предлагаю использовать.
ея нет его адреса. Заранее

Время: 01:21:50

Задание 1 из 1

?

Далее

Инструкции

Чтобы освободить место в папке «Входящие» тебе нужно:

- Удалить рекламные письма с пометкой «Спам»
- Удалить письма, полученные до 15 ноября, касающиеся Нельсоновского проекта, в поле «Тема»
- Сохранить пять самых больших вложений в соответствующие папки на жестком диске ("D:\Вложения")
- Удалить пять писем с самыми большими вложениями, которые ты только что сохранил

Щелкни по кнопке ДАЛЕЕ, когда закончишь выполнение задания.

✉ E-Mail

Файл Правка

Поиск:

Входящие (1)

Удалённые

	От	Тема	Дата	Размер
	Блог Бастер	СПАМ: как вам понравится – 2	19.11.2010 1:07:00	1 KB
	Аллан Ки	СПАМ: лучшие цены	16.11.2010 1:13:00	1 KB
	Тони Майлс	СПАМ: магическая потеря	13.11.2010 16:22:00	1 KB
	Нора Льюис	СПАМ: самое лучшее предлож	08.11.2010 18:19:00	1 KB

От: Блог Бастер

Тема: Спам: Как вам понравится – 2 миллиона ссылок на ваше объявление?!

Размещайте ваши объявления там, где люди смогут прочитать их!

Шаг 1. Сбор свидетельств

Связывает работу, сделанную участниками тестирования со свидетельствами, которые автоматически обсчитываются

Шаг 2. Обобщение свидетельств

Сделать выводы о компетентности участника тестирования на основе полученных свидетельств, обобщив их с помощью матаппарата Байесовских сетей