***ФГБОУ ВО «Кубанский государственный аграрный университет имени И.Т. Трубилина»,***

***Российская Федерация***

**145 Инженирия знаний и интеллектуалные системы. Лабораторная 7. п.4 инструкции 2020-10-19**

**Заголовок**

Лабораторная работа №7: Изучение инструкции по созданию приложений в системе Эйдос и поиск источников данных

**Резюме текста**

Лекция является седьмой лабораторной работой по дисциплине "Инженерия знаний и интеллектуальные системы" для группы БИ1701. Занятие посвящено изучению пункта 4 инструкции по созданию собственного приложения в интеллектуальной системе Эйдос.

**1. Цель и План Занятия:**
Изучить инструкцию по созданию приложений в Эйдос (пункт 4).
Обсудить требования к теме и данным для создания собственного приложения.
Рассмотреть источники данных для машинного обучения.

**2. Требования к Приложению:**
**Новизна:** Приложение не должно повторять уже существующие в системе Эйдос (список доступен на сайте). Допускается взять старую тему и улучшить ее с использованием современных возможностей системы.
**Соответствие нормам:** Тема и данные не должны противоречить морально-этическим нормам и законодательству РФ.
**Источник данных:** Необходимо предоставить ссылку на источник данных и сами исходные данные.

**3. Источники Данных:**
**Kaggle (kaggle.com/datasets):** Рекомендуемый ресурс с большим количеством наборов данных для машинного обучения. Многие наборы связаны с конкурсами, в том числе с призовыми фондами. Есть также данные вне конкурсов для тренировки.
**UCI Machine Learning Repository (archive.ics.uci.edu/ml/index.php):** Классический репозиторий наборов данных, используемый во многих публикациях лектора. Формат данных может быть менее удобен для прямого импорта в Эйдос по сравнению с Kaggle.
**Сайт лектора (professorluka.ru):** Содержит ссылки на Kaggle, UCI и другие ресурсы, а также примеры работ и публикаций с использованием данных из этих источников (например, анализ данных Госстатистики, текстов, изображений).
**Госстатистика (gks.ru):** Данные по регионам РФ, которые можно использовать для анализа взаимосвязей (например, промышленность и экология, качество жизни). Лектор подчеркивает важность корректной интерпретации связей, так как могут существовать скрытые общие причины (например, влияние времени суток, а не показаний часов).
**Другие источники:** Любые доступные наборы данных, соответствующие требованиям.

**4. Форматы Данных и Импорт в Эйдос:**
**Excel (.xls, .xlsx):** Стандартный формат для ввода данных.
**CSV:** Текстовый формат с разделителями (обычно запятая). Лектор упоминает о разработке нового режима в Эйдос (2.3.2.2) для удобного импорта CSV-файлов, который решает проблемы с кодировками (рекомендуется 866 или Windows-1251 для кириллицы) и разделителями (точки вместо запятых), а также позволяет обрабатывать файлы размером более 2 ГБ (лимит DBF), что превышает возможности Excel (лимит около 1 млн строк).
\* **DBF:** Формат баз данных, используемый Эйдос внутренне.

**5. Задание и Дальнейшие Шаги:**
**Найти тему** для собственного приложения.
**Найти источник данных** и сами данные.
З**арегистрироваться** в ResearchGate и РИНЦ (для РИНЦ может потребоваться корпоративная почта, что может быть проблемой для студентов; лектор предлагает помощь с размещением).
**Начать работу** над приложением, следуя инструкции.
**Консультироваться** с преподавателями (по почте или на занятиях).
**Цель:** Создать и защитить собственное приложение, разместить его описание (получив оценку "отлично").

**6. Дополнительные Замечания:**
Лектор упоминает о своем опыте участия в конкурсах Kaggle.
Подчеркивается важность корректной интерпретации выявленных зависимостей, так как корреляция не всегда означает причинно-следственную связь.

**Детальная расшифровка текста**

**Введение**

Здравствуйте, ребята.
Добрый вечер всем.
Добрый вечер, добрый вечер.
Мы не… ну что-то там как-то маловато, что ли? Может, позовёшь остальных там в группе где-нибудь?
Ну, давайте начнём занятие под запись.
Сегодня у нас 19 октября 2020 года. Шестая пара, 17:20-18:50. Седьмая лабораторная работа по группе БИ1701, на которой у нас по плану изучение… изучаем пункт четвёртый инструкции по созданию собственного приложения по дисциплине Инженерия знаний и интеллектуальные системы.
Сейчас я отмечу везде, кто у нас присутствует. И мы начнём занятие.

**1. Изучение инструкции по созданию приложений**

Смотрим, как скачать приложение… эту инструкцию, точнее, извините. Скачать инструкцию. Вот ссылочка вам, ребята, на инструкцию. Где её можно найти, эту инструкцию? Давайте посмотрим.
Если мы выйдем на мой сайт, выберем второй пункт, немножко прокрутим и дойдём до картинки, вот там правее картинки есть задание-инструкция для учащихся по разработке собственного интеллектуального облачного Эйдос-приложения.

**2. Требования к приложению**

**2.1. Новизна**

Требования, ребята, требования. То есть, когда мы выбираем тему приложения, то должны иметь [в виду], что уже довольно много приложений существует, уже создано. Наше приложение не должно повторять тех, которые созданы. Хотя здесь, вот я, хмм, не совсем, так сказать… Добрый вечер… э-э, не на 100% уверен в этом пункте. Дело в том, что некоторые приложения сделаны давно.
Вот посмотрим. Вот нажмёмте по этому пункту. Это список приложений, ребята, уже существующих. И мы видим, что начиная с семнадцатого года, ну, первое приложение просто это лабораторная работа, которую я разместил. А потом идут: семнадцатый год, видите, справа. Потом идут восемнадцатый год, девятнадцатый. И потом уже не очень много идёт приложений двадцатого года.
Я что хочу сказать? Что если взять какое-то старое приложение, которое оформлено не очень хорошо, э-э, не все возможности анализа там использованы, не все там варианты форм выходных, то тогда, в принципе, можно его взять и попробовать просто написать на голову выше. Ну, уже учитывая современный уровень возможностей, которые есть в системе.

**2.2. Соответствие нормам и законодательству**

Это первое требование, чтобы это было новое приложение. Второе. Э-э, исходные данные должны, и сама тема работы не должны противоречить э-э общепринятым в России морально-этическим нормам, а также действующему законодательству Российской Федерации. То есть мы не должны э-э исследовать какие-то вещи, которые, ну, вызывающий характер имеют, провокационный характер и тому подобное. Это нежелательно. То есть мы не будем этого делать.

**2.3. Источник данных**

Вот. Теперь, э-э, тут вот дальше об этом говорится, что вы должны прислать мне источник данных, э-э, и, значит, ссылку на источник данных и исходные данные. Или прислать… показать на занятии.

**2.4. Проверка достоверности модели**

Для того мы должны сделать модель на этих исходных данных и посмотреть, насколько эта модель э-э достоверная. Если она хорошая получается, разумная, вот, то тогда, в общем, ничто нам не мешает, и она оригинальная, то есть нет у нас ещё работ по такой тематике, то нам это вполне для нас приемлемо, и мы можем эту тему выбрать для работы.

**3. Источники данных**

**3.1. Kaggle и UCI**

Где взять исходные данные для своего приложения? Вот здесь вот я вам привожу ссылочки на Kaggle и UCI. Kaggle – это сайт, который содержит большое количество различных данных для машинного обучения. Это источники данных. Очень большое количество различных э-э источников данных.
Что ещё интересно, ребята? Что эти вот данные, которые находятся на сайте Kaggle, они являются данными по конкурсам. И кроме того, что там просто интересные задачи, сами по себе интересные, вот, можно перевести на русский, посмотреть. Кроме этого, э-э, есть ряд конкурсов действующих, и, хмм, есть призовые фонды. То есть, если мы э-э побеждаем, входим в тройку призёров, получив… получивших наилучшие результаты, то там довольно приличные призовые фонды на наш, так сказать, взгляд, на нашу мерку.
Вот. Есть э-э уже прошедшие конкурсы, до… данные, которые тоже можно использовать. И есть э-э ещё такая, я бы сказал, песочница. В общем, данные, которые вне конкурсов, которые можно просто э-э поиграться, потренироваться, сравнить себя с народом, как народ работает.
Вот. И их там, я не знаю, даже как сказать. Ну, в общем, огромное количество этих данных. Просто огромное.
Я когда-то на Kaggle зарегистрировался довольно давно. Вот, и участвовал в нескольких задачах. Даже попробовал несколько задач решить.
Вот. Ну, в общем, э-э, не особо так получилось. Э-э, ну, как сказать? Значит, в одной задаче э-э получилось так, что мои решения оказались лучшими, чем у 85% участников, а в другом случае лучшими, чем у 91% участников. Если я не ошибаюсь. Вот здесь вот об этом мы видим информацию.
Вот. Ну я э-э специально не занимался тем, чтобы занять призовое место. Это требует времени, требуется прямо э-э уделять много этому времени, чтобы получить какой-то хороший результат.
Вот. У меня этого времени нет, поэтому, в общем, я вам… есть то, что есть. Вот. Ну, в принципе, можно было бы попробовать решить какую-то задачу на более высоком уровне, э-э, приближающемся к призовым местам. Но для этого требуется там уже э-э усилия предпринять определённые.
Скажу вам, что я э-э посмотрел потом, какие есть пользователи в Kaggle, какие у них успехи у наиболее продвинутых. И могу вам сказать, что есть такие, ну я не знаю, как это сформулировать так без жаргона. Ну, гиганты какие-то, наверное, так можно сказать. Просто гиганты, которые э-э постоянно занимают призовые места.
И недавно, ну как недавно, года два, наверное, назад произошла сенсация такая определённая, шум возник в этой связи на этом же Kaggle, связанный с тем, что э-э появился какой-то молодой э-э парень, ну насколько молодой? Ну он очень молодой. То есть он учащийся старших классов, школьник, который вдруг стал занимать регулярно сразу в нескольких чемпионатах высокие, то есть призовые места, первые места. То есть какой-то такой гений появился молодой, который даже ещё не имел высшего образования, ну такого уровня высокого, что он смог э-э занять несколько призовых мест.
А призовые фонды, они такие там 5.000, там 10.000, там 30.000, 100.000 я видел. Смотря какая задача, кто её там выставил. И э-э тех людей, кто занимают призовые места, особенно регулярно если, очень большая вероятность, что их пригласят на работу какие-то ведущие фирмы. То есть это вот э-э практика такая.

**3.2. Сайт лектора и Госстатистика**

Вот. Есть ещё… так, ну вот я вам послал сейчас вот ссылочки на некоторые статьи. Ещё есть, сейчас ещё пошлю, наверное. До последней дойдём. Вот это, по-моему, уже самая последняя.
Вот. То есть я вам привожу ряд примеров, статей, которые используют э-э данные сайта UCI. Вот вы можете взять эти статьи и наподобие что-то э-э описать… сделать описание наподобие этих статей в своих работах.
Но здесь я вам могу сказать что интересного? Что я всегда в статьях указываю источник данных обязательно, ссылаюсь на авторов этого источника данных, даю ссылочку. Вот. Потом идёт описание уже. Использование этих данных для создания приложений для машинного обучения.
(Примеры: ветеринарный тест, атрибуция текстов, анализ президентских обращений).
То есть можно довольно интересные работы сделать.
Вот. И даже это сделал этот текст э-э подписи, ссылочка на него сделал подписи на свои письма. Все об этом рассказываю.
Вот один из студентов мне уже прислал данные. Я их посмотрю и скажу ему, берём мы эту тему или нет.
Ну, в общем, можно сказать так, что вкратце я вам рассказал, как мы должны дальше действовать, что вы должны делать. Вот. И мы с Георгием Александровичем будем вам помогать консультациями. И утверждаем тему. И после этого вы начинаете делать описание этого приложения по шаблону. Сейчас здесь приведено четыре шаблона. Вот. Ну как шаблоны описания можно рассматривать все мои статьи практически, какие у меня есть, э-э связанные с решением конкретных задач. Есть статьи теоретические, а есть такие, которые можно как э-э примеры лабораторных работ рассматривать или каких-то научных исследований.
Вот я сейчас вам показал, где эти статьи есть. Они есть на той же страничке, на которой можно скачать систему Эйдос, на той же страничке, где можно скачать обновление.

**4. Форматы данных и импорт в Эйдос**

Вот. То есть вы можете использовать также и не только экселевский файл, но и CSV файл. Потому что сегодня я сделал режим, закончил отладку режима, который это позволяет делать. Сегодня размещу новую версию, по-видимому.
(Объяснение режима CSV DBF конвертера, кодировок 866/1251, разделителей, лимитов Excel и DBF).
То есть мы можем таким образом вводить данные из CSV файла с сегодняшнего дня. Совершенно спокойно. Не используя никаких конвертеров, которые надо ещё искать, неизвестно как, насколько мы найдём то или не то, платные или бесплатные. И, в общем, мы вот ввели данные. Дальше можем уже осуществлять синтез модели и полностью работать.
И ещё я сделал обновление в алгоритме… в развитом алгоритме принятия решений вот в этом режиме. И в режиме… порядок преобразования данных в информацию, её в знание. Тоже там сделал обновление. То есть внёс изменения в эти файлы, которые, на мой взгляд, улучшают возможности изложения материала учебного.

**4.1. Пример анализа данных Госстатистики**

Вот. Ну, в общем, я уже знаете, что система Эйдос выявляет зависимости между факторами какими-то и последствиями их влияния. А в статистике есть такая особенность, ребята. Сейчас вы слушайте внимательно. Вся эта статистика привязана к регионам Российской Федерации. То есть вот слева, вот здесь вот шапочка, Алтайский край там и так далее. Сейчас я вот сделаю пошире, чтобы видно было. Видно? Регионы России. И потом по алфавиту, видите, идёт?
Так вот все эти таблицы статистические, вот эту левую часть имеют одинаковую. Это что означает? Что мы можем взять таблицу какую-то, ну, допустим, развитие промышленности, тяжёлой промышленности, там или средней промышленности, и, допустим, экологические параметры, вот эти же самые, которые здесь вот у меня. И вы узнаете, как развитие промышленности влияет на экологию. А можете взять развитие промышленности и сразу вот эти вот параметры качества жизни. Узнаете, как влияют эти параметры на качество жизни. Можете взять, допустим, развитие сети 4G, 5G, да? Правильно? И посмотреть, как это связано со статистикой там заболеваний определённого рода или и смертности от них. Тот же самый COVID, о котором вот сейчас вот говорят все о нём. И возникает вопрос такой: а это что говорят, оно как-то оправдано, обосновано или не очень? Ну вы можете просто это узнать, оправдано это, обосновано или не очень. То есть как-то связано это одно с другим или нет?
Я так думаю, что моё мнение такое насчёт того, что как влияет ли сам 5G, я не знаю. Не исключено, что и влияет, я не исключаю такой особенности, возможности. Но я думаю, что здесь больше связано, знаете с чем? 5G где появляется? А там, где высокая активность населения, там, где высокая активность, высокий уровень развития технологий, коммуникаций, да? Вот. А там, где такой уровень развития технологий и коммуникаций, там высокий уровень развития организационных связей различных. То есть там люди мобильны, люди много контактируют друг с другом. То есть получается так, что не может быть, это даже не является причиной этот самый этот 5G. Но он появляется там, где у людей в их образе жизни есть такие особенности, которые способствуют распространению ковида. Вот и всё, понимаете? То есть это как сказать?

**4.2. Корреляция и причинность (пример с часами)**

Ну если мы посмотрим на то, как ведут себя люди в различное время суток. Утром они просыпаются, моются, завтракают и начинают работать. А вечером они ложатся спать, а ночью спят. И посмотрим на показания часов, то получится, что есть очень высокая степень взаимосвязи между показаниями часов и тем, что делают люди. Отсюда можно сделать вывод такой, что показания часов обуславливают поведение людей. Вот всё вроде как бы вот так более-менее логично, так и казалось бы, и даже похоже на правду. Но вот последний вывод какой-то очень странный. Вам не кажется, нет, ребята? Вот я сейчас хочу, чтобы вы ответили, сказали: кажется вам странным это или не кажется?
Ну да.
А вы понимаете, что показания часов не влияют на поведение людей? Разве что если только не брать сказку про э-э как же она называлась? Аленький цветочек, вот вспомнил, Аленький цветочек. Давно это дело было. Вот там сестрёночки благодетельницы перевели часы и закрыли ставни. И она не смогла вовремя вернуться, и чудище там чуть не умерло. Ну не тогда, то есть можно сделать так, что и показания часов будут влиять на жизнь. Вот. Но вообще-то, конечно, не сами показания часов влияют, а солнце, правильно? Время суток точнее даже просто. Вот. И фактически о чём идёт речь? О том, что время суток влияет на наше поведение. И время суток влияет на показания часов. Поэтому если мы вычисляем зависимость поведения людей от показаний часов, то у нас получается, что высокая степень зависимости. Но на самом деле эта зависимость не от показаний часов, а от времени суток.

**4.3. Сложность интерпретации связей**

Вот. Но этого мы можем не понимать. Вот что интересно. То есть когда про часы я привёл пример, то там всё, в общем, более-менее понятно, что это так. А вот если там будет более сложная ситуация, вот, допустим, производство, развитие производства тяжёлого, скажем, тяжёлой промышленности, экологические показатели, охрана окружающей среды и определённые виды заболеваемости и качества жизни. То нам может показаться, что действительно развитие производства влияет на это. Поняли, да, мою мысль? Но на самом деле на это влияет не развитие производства. Развитие производства является, скажем так, технологической средой. А среда имеет несколько уровней иерархической структуры своей. То есть это природная среда, технологическая среда, организационная среда, включающая взаимосвязи людей с помощью потоков вещественных, предметов, жидкостей, газов, с помощью передачи энергии и с помощью передачи информации, информационная среда, которая сейчас очевидно, она выходит на первый план и обуславливает и остальные. Вот. А также выше идёт уже та организационная среда, а выше идёт экономическая среда. То есть это отношения людей, связанные со стоимостью, финансами, финансовыми потоками, в частности, и генерацией финансов, там, тратами, передачей. Вот. И потом дальше идёт уже политическая структура. А ещё выше вот эта культурная, этнографическая уже среда, культурная этносреда и цели, ценности, мотивации и духовные ценности народа определённого, включая религиозные ценности. И такие как вот, скажем, декларация независимости США там провозглашают, связанные, ну, со светским обществом, но фундаментальные. И вот смотрим мы на это всё и понимаем, что вообще-то, конечно, все эти виды среды влияют на то, как люди живут в этом обществе, где такая окружающая среда со всеми этими структурными уровнями организации. Вот. Ну то есть не само развитие технологии влияет, а всё, что с этим связано: и образ жизни, и образ взаимодействий и так далее, и так далее, и так далее. То есть всё несколько сложнее, чем кажется на первый взгляд. Но формально можно посмотреть и сказать: "Ага, вот это вот у нас, если развито технологическая среда, тогда вот это и вот это будет в плане, скажем, заболеваемости или каких-то критериев качества жизни". Ну, в общем, это будет правильно. Хотя интерпретировать это содержательно будет довольно проблематично. Ну, когда вот, допустим, мы говорим про стрелки часов, как они влияют на жизнь, то если мы начнём интерпретировать это содержательно, почему стрелки часов влияют на то, что люди делают, то довольно трудно содержательно это объяснить. Вот. То есть влияние есть, но оно какое-то, наверное, всё-таки информационное это влияние. То есть люди получают определённую информацию о чём-то, что их интересует, и что для них существенно. И эта информация уже влияет на их поведение. Эту информацию они могут получать не только от часов, они могут её на телефоне посмотреть время, тоже ж точно так же это будет влиять или на компьютере. Вот. И также влияет и другая информация, не только о времени суток, а любая информация, которая поступает к нам по различным каналам связи: мобильной связи и интернета, конечно, Ватсаппа, всё прочее. То есть всё, что мы узнаём, всё это существенно влияет на наше поведение. Ну, скажем так, не всё существенно, а так более или менее существенно в том или ином направлении влияет. Вот если включишь телевизор, посмотришь передачу, то возникает такое ощущение, что вообще там конец света. Это влияет на наше поведение, на наше настроение, на наши цели, ценности? Ну, конечно, влияет, ребята. Вот мы можем выявить зависимости о том, как вот влияет то, что там показывают на число там каких-то правопротивоправных действий, например. Я вам скажу, что, в общем-то, люди привыкают считать, что а, это, наверное, нормально, взять там машину перевернуть там или ещё что-нибудь такое сделать. Потому что мы это видим каждый день много раз в разных странах. Слава Богу, у нас мы этого не видим, но, как говорится, люди, если к этому привыкнут, и будут считать, что это как бы само собой разумеется, переворачивать машины, когда им что-то не нравится, ну тогда и у нас увидим. Всё это можно исследовать.

**5. Задание и дальнейшие шаги**

Так, теперь, пожалуйста, какие вопросы у вас по ходу этих работ? То есть я сейчас вам даю задание, э-э, начать самостоятельно выполнять пункты четыре и дальше в контакте с нами. Контакт с нами, значит, здесь почта указана. Вот. Даже, по-моему, я и Skype показывал. Ну, Skype нет необходимости, вот почтой пользуйтесь вот этой. Вот. Можете что-то спрашивать. И на занятиях, занятия наши будут теперь этому посвящены. И по почте спрашивайте. Буду находить время и по почте тоже вам отвечать.
Вопросов на данный момент нет.
Ну тогда надо искать тему. Вот. Порыться там на этих сайтах. И параллельно регистрироваться в ResearchGate и РИНЦ. Для регистрации... в общем, для регистрации в ResearchGate необходим корпоративный адрес электронной почты. Вот с этим может быть проблема. Насколько я понимаю, потому что в нашем университете почему-то студентам не присваивается корпоративный адрес электронной почты. Непонятно, э-э, почему. Потому что в других университетах, где я работаю, это уже несколько лет так присваивается. У нас до сих пор, значит, этого нет. А это необходимо, потому что вам нужно идентифицироваться там как учащимся вуза, для того, чтобы вас там зарегистрировали. Вот. Ну этих ребят прикрыть вообще ничем невозможно, которые проверяют. Они найдут в этих работах массу недостатков и скажут, что этот автор вообще ничего не соображает, и вот бесполезно пытаться их чем-то прикрыть там. Ну хотя есть кое-что, чем можно прикрыться, но я не буду конкретизировать.

**Заключение**

Ребята, давайте я сейчас, пожалуйста, вопросы, осталось 10 минут. Ну Данила сказал, что нет вопросов. Ну хорошо. А у других тоже нет, что ли? Значит, давайте тогда к следующему занятию договоримся, чтобы у вас были темы и данные. А я вам пообещаю, что сегодня я постараюсь разместить, ну правда поздно вечером, у меня занятия до 21:30 идут. Поэтому я поздно вечером, э-э, ну, к полуночи, размещу новую версию системы, которая будет и CSV файлы вводить. А, вот что я вам рассказывал, что я проводил эксперимент при отладке этих режимов на 4 млн наблюдений, которые описаны в CSV файле. Всё около тридцати этих полей. То есть это в Excel не влезает в четыре раза. Понимаете, только четверть в Excel влезает. При этом получился файл исходных данных около 760 МБ. Ну примерно в три раза, чем 2 ГБ меньше. То есть меньше, чем 3... чем 2 ГБ в три раза. Почему я говорю про 2 ГБ? Потому что это максимальный размер файла исходных данных стандарта dBASE, в котором есть... есть пред... предконвертер, ну то есть, то есть пред-интерфейс, э-э, CSV, DBF конвертер. Я его буквально вот вчера, сегодня реализовал на этом языке программирования, на котором написана система. Вот. И там, э-э, возможность этого файла 2 ГБ максимальная. То есть максимальный его размер. Значит, исходя из того, что вот я сейчас увидел, что он на 4 млн 765 МБ занимает, я предположил, что, в принципе, может быть там от 10 до 12 млн наблюдений можно обрабатывать таким образом. Ну то есть система Эйдос, она, как скажем, довольно большие объёмы исходных данных может обрабатывать. Чисто технически, около 2 млн... то есть 12 млн наблюдений. Ну это так неплохо, в принципе. Я таких задач не решал. У меня самое большее 4 млн записей решал задачи. Ну я знаю точно, что во многих системах, э-э, в которых пытаются люди обрабатывать такие объёмы данных, это сделать не удаётся. То есть системы не позволяют, у них есть ограничения свои. У системы Эйдос тоже есть свои ограничения. Я когда её разрабатывал, мне в голову не могло прийти, что могут быть миллионы наблюдений, например. То есть я вообще такого кошмарного сна... мне даже фантазии не хватало это представить. Я могу вам сказать, что у меня на компьютере было 200... 200 МБ Винчестер стоял, когда я её разрабатывал эту систему, в девяносто втором году, девяносто третьем. Вот такие были объёмы Винчестеров. Ну 200, там 180 МБ, 220 МБ. И я знал, что язык программирования обеспечивает обработку баз данных размером 2 ГБ. И мне казалось, что этого достаточно. Понимаете? То есть, ну куда, зачем мне нужно больше, чем 2 ГБ, если у меня Винчестер 180 МБ? Понимаете, да, о чём речь, ребята? Мне казалось, что этого вполне достаточно. Я уверен, что это казалось и разработчикам языка программирования. Они тоже думали, что 2 ГБ база данных достаточно, потому что таких Винчестеров нет. Самое большее 800 МБ, которые тогда существовали вообще по литературе. Ну и всё. Значит, тогда вопрос этот не возникал, что это маловато. А сейчас у меня очень острое ощущение, что 2 ГБ - это мало, понимаете? Кажется, что ну вообще. То есть, ну, многие задачи, конечно, помещаются в эти размеры, в эти размерности. Но есть задачи, которые, ну, не получится решить при таких ограничениях на размеры баз данных. Поэтому я в системе Эйдос предусмотрел ряд ухищрений таких, там есть там свои хитрости, моменты, которые позволяют и при таких ограничениях решать задачи большей размерности. Вот. Ну вот по исходным данным я сказал, около 12 млн. Ну там, почему я говорю около, а не точно? Потому что от числа полей зависит. Если полей будет, например, 10, то тогда, значит, 20 млн будет записей можно обработать наблюдений. Если полей будет, например, 100, ну тогда можно будет там, ну, скажем, 5-6 млн записей наблюдений. Ну примерно так. То есть вам смысл понятен. Ещё вопросы, пожалуйста, осталось 5 минут. Я отвечу на один-два вопроса, хватит ещё.