Министерство науки и высшего образования РФ

Федеральное государственное автономное

образовательное учреждение высшего образования

«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт космических и информационных технологий

институт

Кафедра прикладной информатики

кафедра

**ОТЧЕТ О ПРАКТИЧЕСКОЙ РАБОТЕ № 6**

Интеллектуальный анализ данных

наименование дисциплины

Методы извлечения знаний (вариант 9)

тема (вариант)

Преподаватель \_\_\_\_\_\_\_\_\_\_ Т.Н. Сизова

подпись, дата инициалы, фамилия

Студент КИ22-21б 032215739 \_\_\_\_\_\_\_\_\_\_ А.С. Полешко

номер группы номер зачетной книжки подпись, дата инициалы, фамилия

Красноярск 2025

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 3](#_Toc196492704)

[Ход работы 4](#_Toc196492705)

[Часть 1. 4](#_Toc196492706)

[Часть 2. 8](#_Toc196492707)

[ЗАКЛЮЧЕНИЕ 10](#_Toc196492708)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 11](#_Toc196492709)

# **ВВЕДЕНИЕ**

Цель работы:рассмотрение процесса получения и структурирования знаний являющийся при создании базы знаний и представлена классификация методов извлечения знаний.

# **Ход работы**

**Задача:**

1. Что такое извлечение знаний? На какие этапы оно делится?

2. Каковы функции инженера по знаниям и эксперта? С какими трудностями приходится сталкиваться в процессе вербализации знаний?

3. Охарактеризуйте существующие стратегии получения знаний. В чем их принципиальная разница?

4. Объясните суть теоретических аспектов извлечения знаний. Как вы думаете, необходимо ли было их рассмотрение в данной главе? Почему?

5. Попытайтесь самостоятельно разработать примеры получения знаний для каждого из методов, приведенных на рис. 15.1

**Решение:**

**1. Что такое извлечение знаний? На какие этапы оно делится?**

Извлечение знаний - процесс взаимодействия инженера по знаниям с источником знаний (экспертом), в результате которого становятся явными процесс рассуждений специалистов при принятии решения и структура их представлений о предметной области.

Обычно извлечение знаний включает следующие этапы: выбор данных (отбор), очистка (предобработка), трансформация, моделирование и интерпретация результатов.

**2. Каковы функции инженера по знаниям и эксперта? С какими трудностями приходится сталкиваться в процессе вербализации знаний?**

Эксперт – высококвалифицированный специалист предметной области, согласившийся поделиться своим опытом. Эксперт обладает глубокими знаниями в определенной области и предоставляет экспертную оценку и контекст в процессе извлечения знаний. Эксперты часто участвуют в валидации полученных знаний и помогают установить правильный контекст для их применения.

Инженер по знаниям (когнитолог, аналитик, инженер - интерпретатор) – специалист в области искусственного интеллекта, выступающий промежуточным звеном между экспертом и базой знаний. Инженер по знаниям занимается структурированием, представлением и использованием знаний в интеллектуальных системах. Он определяет цели сбора данных, разрабатывает модели знаний, создает алгоритмы для обработки данных и обеспечивает надежное хранение знаний.

Трудности в выражении знаний: экспертам бывает сложно вербализировать свои знания, особенно если они являются интуитивными или неформализованными.

Проблемы с пониманием контекста: инженерам по знаниям может быть трудно полностью понять контекст и специфику знаний экспертов.

Различия в терминологии: возможны недоразумения из-за использования разных терминов и понятий.

Когнитивные искажения: эксперты могут бессознательно упускать важные детали или придавать слишком большое значение незначительным аспектам.

**3. Охарактеризуйте существующие стратегии получения знаний. В чем их принципиальная разница?**

Существует несколько стратегий получения знаний. Наиболее распространенные:

* Извлечение знаний – это длительная и трудоемкая процедура, когда когнитологу необходимо воссоздать модель предметной области, которой пользуются эксперты для принятия решения.
* Приобретение знаний представляет собой процесс наполнения БЗ экспертом с использованием специализированных программных средств при прямом контакте с ЭС.
* Формирование знаний является процессом анализа данных и выявления скрытых закономерностей с использованием специального математического аппарата и программных средств.

Основное различие между этими стратегиями заключается в подходах к получению знаний и источниках информации, на которых они основаны.

**4. Объясните суть теоретических аспектов извлечения знаний. Как вы думаете, необходимо ли было их рассмотрение в данной главе? Почему?**

Теоретические аспекты извлечения знаний делятся на:

* Психологический аспект извлечения знаний;
* Лингвистический аспект извлечения знаний;
* Гносеологический аспект извлечения знаний.

Из трех аспектов извлечения знаний психологический – является ведущим, поскольку он определяет успешность и эффективность взаимодействия инженера по знаниям (аналитика) с основным источником знаний – экспертом-профессионалом. Психологический аспект выделяется еще и потому, что извлечение знаний происходит чаще всего в процессе непосредственного общения разработчиков системы. А в общении психология является доминантной.

Лингвистический аспект касается исследований языковых проблем, так как язык — это основное средство общения в процессе извлечения знаний.

Гносеология — это раздел философии, связанный с теорией познания, или теорией отражения действительности в сознании человека. Гносеологический аспект извлечения знаний объединяет методологические проблемы получения нового научного знания, поскольку при создании БЗ эксперт часто впервые формулирует некоторые закономерности, до того момента составлявшие его личный опыт.

Да, рассмотрение теоретических аспектов извлечения знаний является необходимым для понимания всех аспектов процесса извлечения знаний.

**5. Попытайтесь самостоятельно разработать примеры получения знаний для каждого из методов, приведенных на рис. 1.**

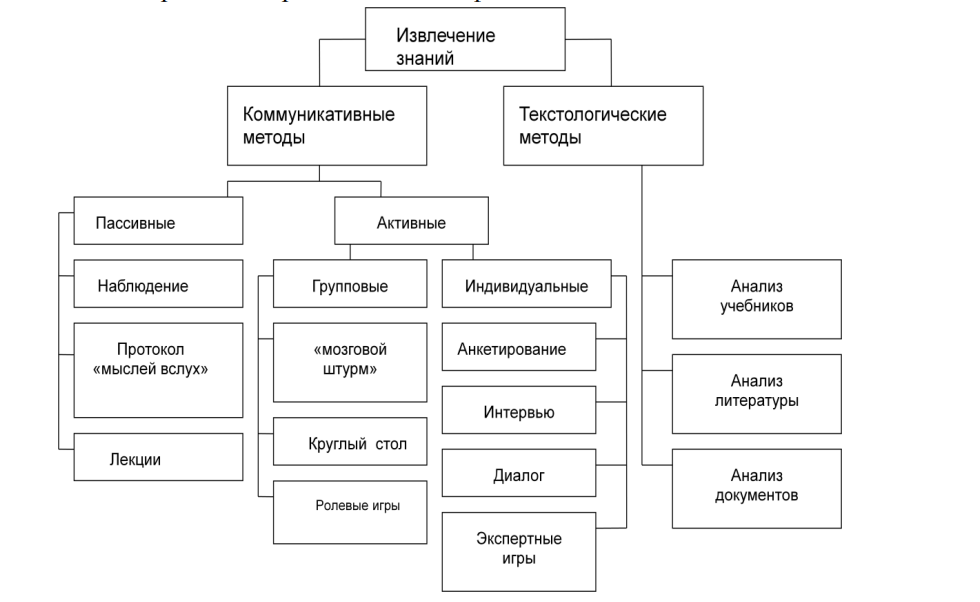


Рисунок – Классификация методов извлечения знаний

1. Коммуникативные методы

1.1 Пассивные

– *Наблюдение*: наблюдение за работой шеф-повара в ресторане, чтобы узнать о процессе приготовления блюд.

– *Протокол «мыслей вслух»*: столяр вслух комментирует каждое своё действие, объясняет причины своих решений, делится опытом и личными принципами.

– *Лекции*: посещение лекций по интеллектуальному анализу данных.

1.2 Активные

1.2.1 Индивидуальные:

– *Анкетирование*: проведение опроса среди студентов, чтобы узнать их предпочтения по учебным предметам.

– *Интервью*: интервью с поваром о правилах сочетаемости продуктов для кулинарной экспертной системы.

– *Диалог*: беседа с преподавателем по теме «Формулы Байеса».

– *Экспертные игры*: игра с врачом, в которой инженер описывает симптомы болезни, чтобы выявить последовательность диагностических шагов, которые эксперт не может вербализовать напрямую.

1.2.2 Групповые:

– *Мозговой штурм*: сессия с врачами: «Какие неочевидные симптомы указывают на раннюю стадию панкреатита?». Каждый выдвигает свою идею.

– *Круглый стол*: обсуждение спорных диагнозов с участием гастроэнтеролога, хирурга и терапевта.

– *Ролевые игры*: суд по медицинской ошибке – «Пациенту назначили препарат Х, но развились осложнения. Была ли ошибка?». Участвуют врачи, адвокаты, судья.

2 Текстологические методы

– *Анализ учебников*: чтение учебника по математике для подготовки к контрольной работе.

– *Анализ литературы*: изучение руководства пользователя по эксплуатации кондиционера.

– *Анализ документов*: чтение условий открытия вклада в банке.

# **ЗАКЛЮЧЕНИЕ**

В ходе выполнения работы были получены знания, умения и навыки по созданию базы знаний, представляющей собой нечеткие модели знаний.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Электронный курс образовательной программы по дисциплине «Интеллектуальный анализ данных» // Система электронного обучения Сибирского Федерального университета: [сайт]. – Красноярск, 2010 – . – URL: <https://e.sfu-kras.ru/course/view.php?id=1109> (дата обращения: 25.04.2025). – Режим доступа: для авториз. пользователей. – Текст: электронный.