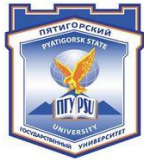




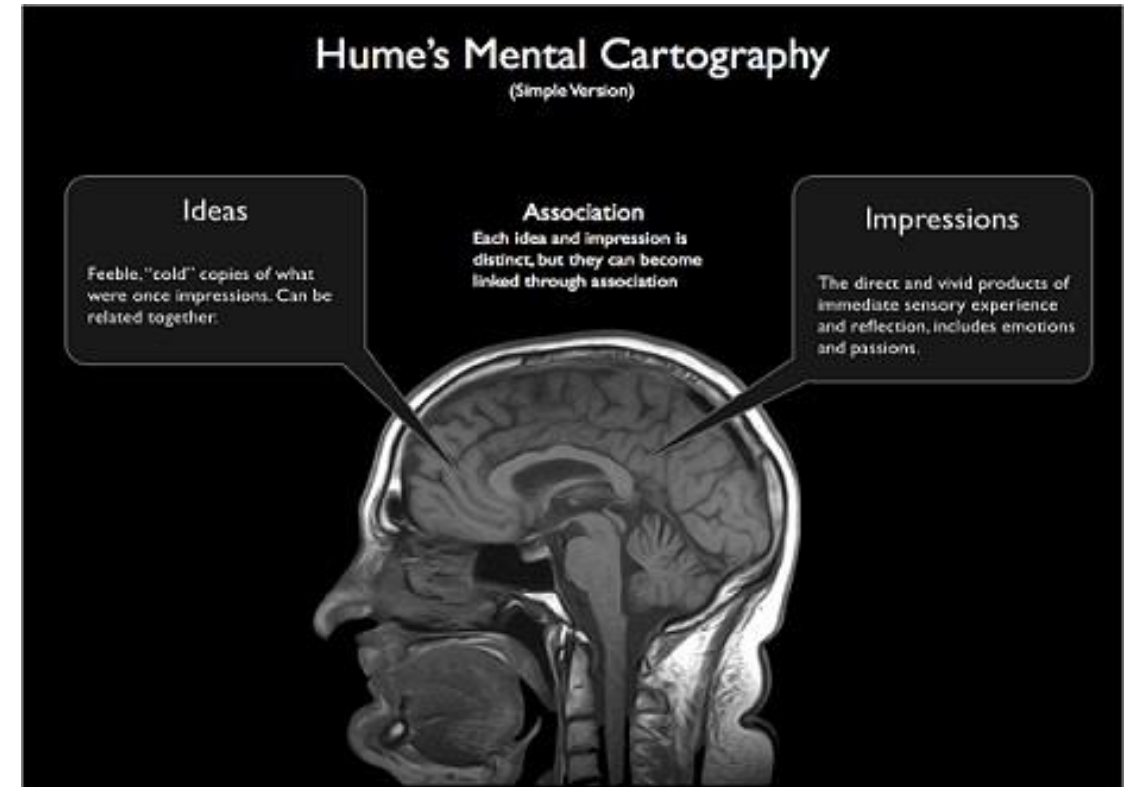
**Машинное «знать», «уметь», «владеть»:
семантика данных и представление ментального
содержания**



План доклада

- 1. Введение в проблему
- 2. Пять аналогий интеграции данных
- 3. Анализ локативных конструкций

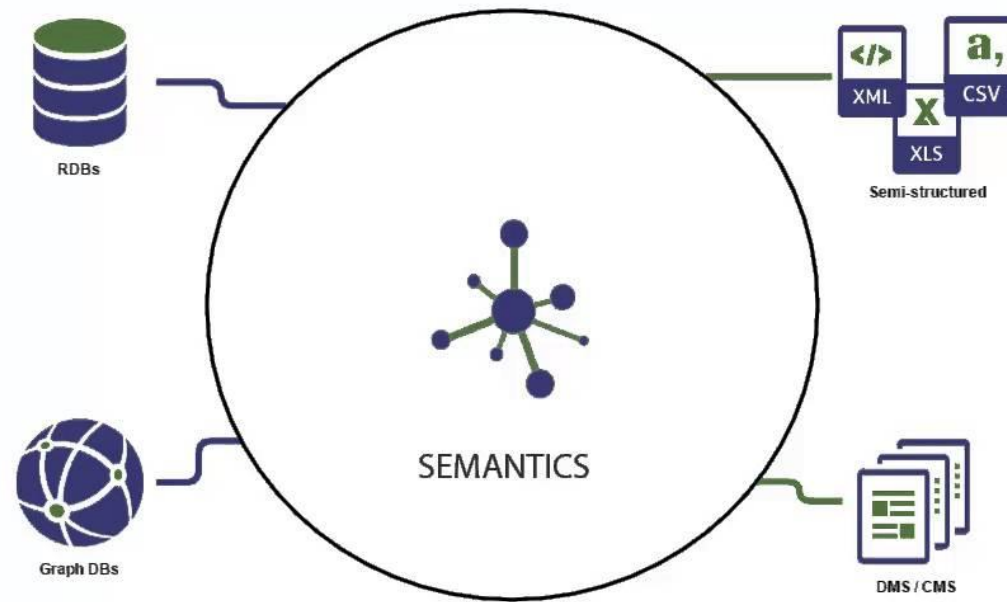
- В первую очередь необходимо отличать понятие «данные» из области компьютерных наук от когнитивных понятий типа «знание», «значение», «смысл», которые в определенном контексте тоже понимаются как данные.



Ключевое отличие представленных ПОНЯТИЙ СОСТОИТ В СЛЕДУЮЩЕМ:

- **данные в искусственных системах** – это способ представления информации в формализованном виде, пригодном для хранения, передачи и обработки;
- **данные в общем виде** (например, база знаний) – представление фактов, понятий или инструкций в форме, приемлемой для общения, интерпретации или обработки человеком или с помощью автоматических средств.

Knowledge Graph



- С развитием семантического подхода в компьютерных моделях управления знаниями происходит функциональное раздвоение:
- данные (синтаксис) – в машине;
- контекст (семантика), задающий критерии релевантности данных, – в человеке.

Introduction TO Syntax & Semantics



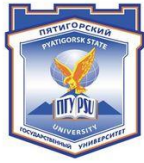
Syntax:-

Like Ordinary language English, programming Languages have Syntax.

The Syntax of a (programming) Language is a set of rules that defines what sequences of **Symbols** are considered to be valid expressions (programs) in the language.

OR

The Syntax of a programming Language is what the program



- Основной тезис:
- между формально-синтаксическим и семантическим способами интеграции данных существуют отличия, подобные отличиям между компьютерными и концептуально-семантическими моделями сознания.

Функции интеграции данных

- Основная задача технологии интеграции данных состоит в преодолении многочисленных проявлений неоднородности и неоднозначности.
- «Системы имеют разную функциональность, используют различные типы данных (алфавитно-цифровые и медийные, структурированные и неструктурированные), их компоненты различаются по автономности, имеют различную производительность».



Интересные аналогии

- Common Data Storage
- Integration by Middleware
- Integration by Applications
- Common User Interface
- Manual Integration

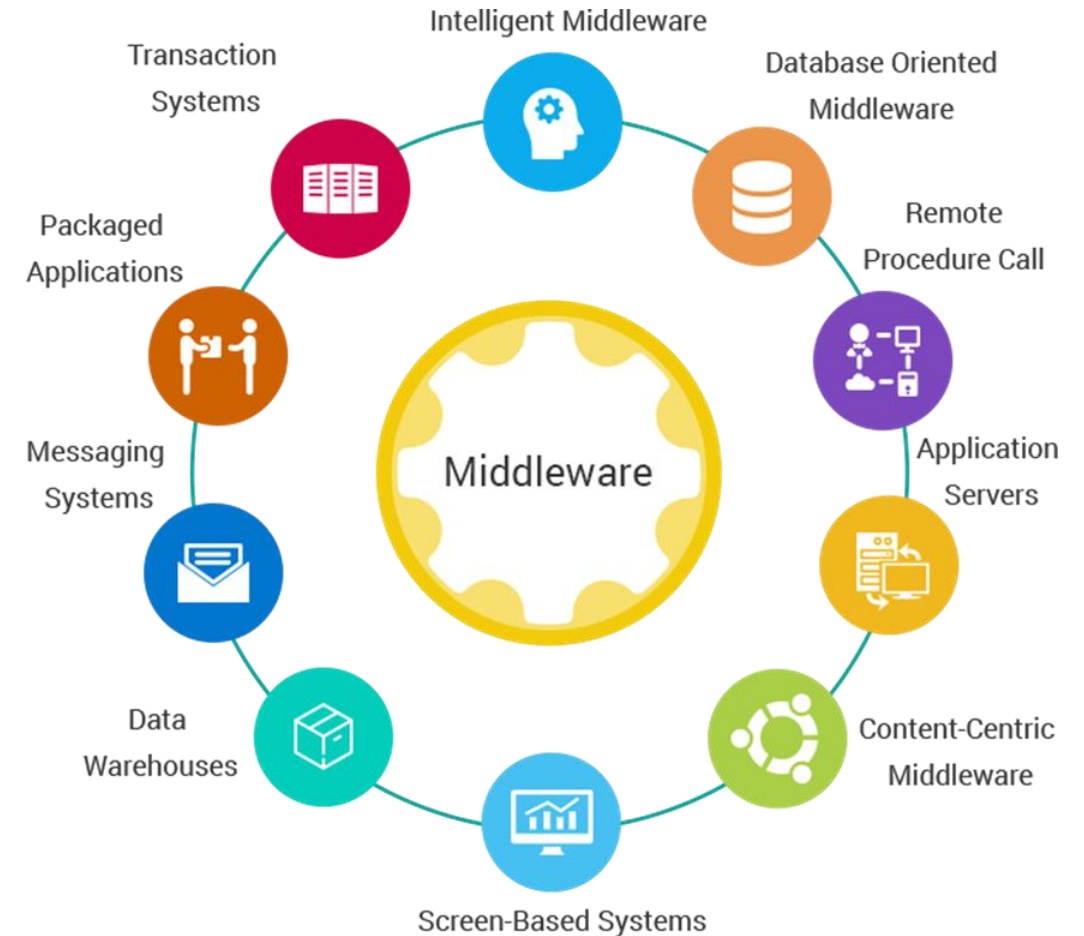
Аналогия 1:

- Под *Общими системами хранения* (Common Data Storage) можно понимать материальный субстрат мозга в совокупности с высшими психическими уровнями (сознание), чья структурная организация позволяет хранить и обрабатывать информацию в таком виде, чтобы был возможен переход на следующие уровни.



Аналогия 2:

- *Интеграция средствами промежуточного слоя (Integration by Middleware)* может являться аналогом фодоровского Языка мысли (Mentalese), который представляет собой операциональную систему абстрактных символов, формирование элементов пропозиционального знания.



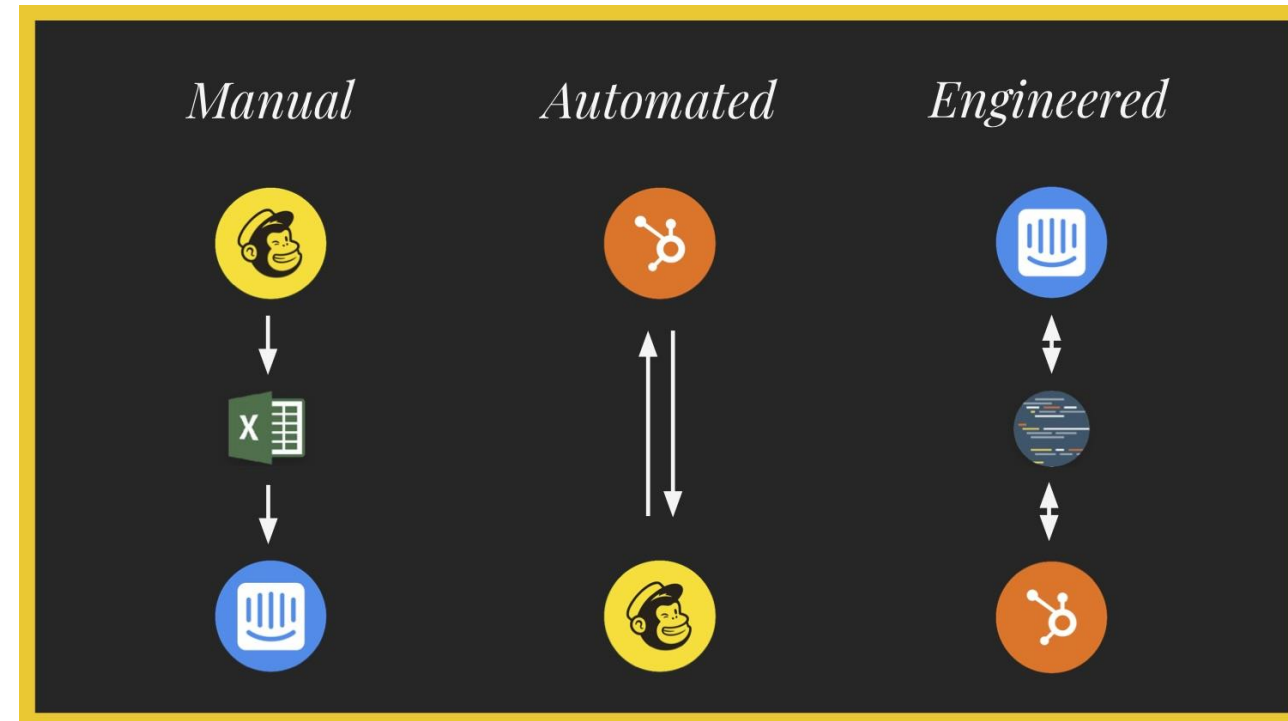
Аналогия 3:

- *Интеграция средствами приложения* (Integration by Applications). В терминах интеграции данных приложения «обеспечивают доступ к разным источникам данных и возвращают пользователю обобщенные результаты» [Черняк, 2013].
- В традиционных теориях сознания обобщение когнитивных процедур сводится к квалиа и ментальным репрезентациям.



Аналогия 5:

- *Интеграция вручную* (Manual Integration) происходит тогда, когда пользователь сам объединяет данные, применяя различные типы интерфейсов и языки запросов.
- Сознательное (интенциональное) производство смысла (речь, создание текстов, коммуникативное поведение) представляется как создание метаязыковых конструкций, интегрирующих пропозициональное содержание с общим внепропозициональным знанием.



- Информационные процессы интеграции данных отчасти объясняют принципы семантических представлений знаний в различных контекстуальных системах.
- Ключевое отличие «слепого» вычисления релевантных значений от ментального процесса семиозиса состоит в том, что основанием для вывода в первом случае является иерархия классов, протоколы правил и применяемые к ним эвристики и стратегии,
- во втором случае основанием для формирования знания является качественный опыт и метафорическая концептуализация, реализуемая в коммуникативной практике.

Сегодня эти высказывания выглядят несколько устаревшими

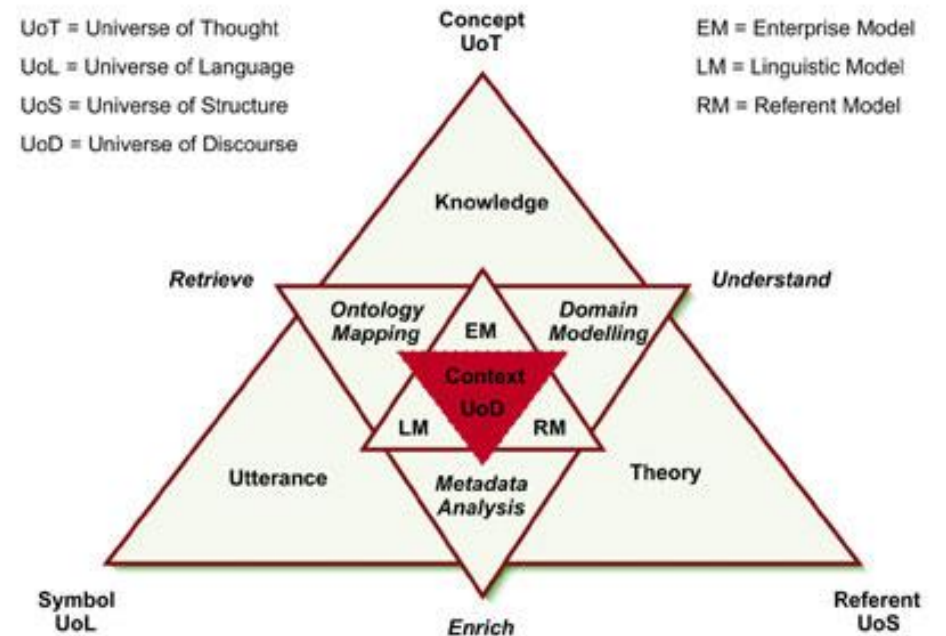


Язык как окно в человеческую природу

- Особое место в исследованиях по проблеме концептуализации пространственных отношений занимают работы по анализу семантики проективных высказываний.
- Языковые структуры становятся своеобразными «свидетельствами» о свойствах когнитивных процессов. Отношения между локативными объектами («Что находится?») и референциальными объектами («Относительно чего находится?») фиксируют навигационные метки.

Концептуальная «разметка» в естественном языке

- Значение сводится к концептуализации.
- Полисемия является нормой и представлена в «пучках» категориальных семантических связей.
- Существуют базовые пространственные примитивы (время, расстояние, цвет, текстура), на основании которых строятся вариативные концепты.
- Семантические структуры выводятся из положения «профилей» относительно базовых значений.
- Семантические структуры включают в себя конвенциональные группы, в которых одно и то же содержание выводится различными способами [Olivier; Ananiadou; Maeda; Tsujii, 1992].



Пример из других естественных сигнальных систем

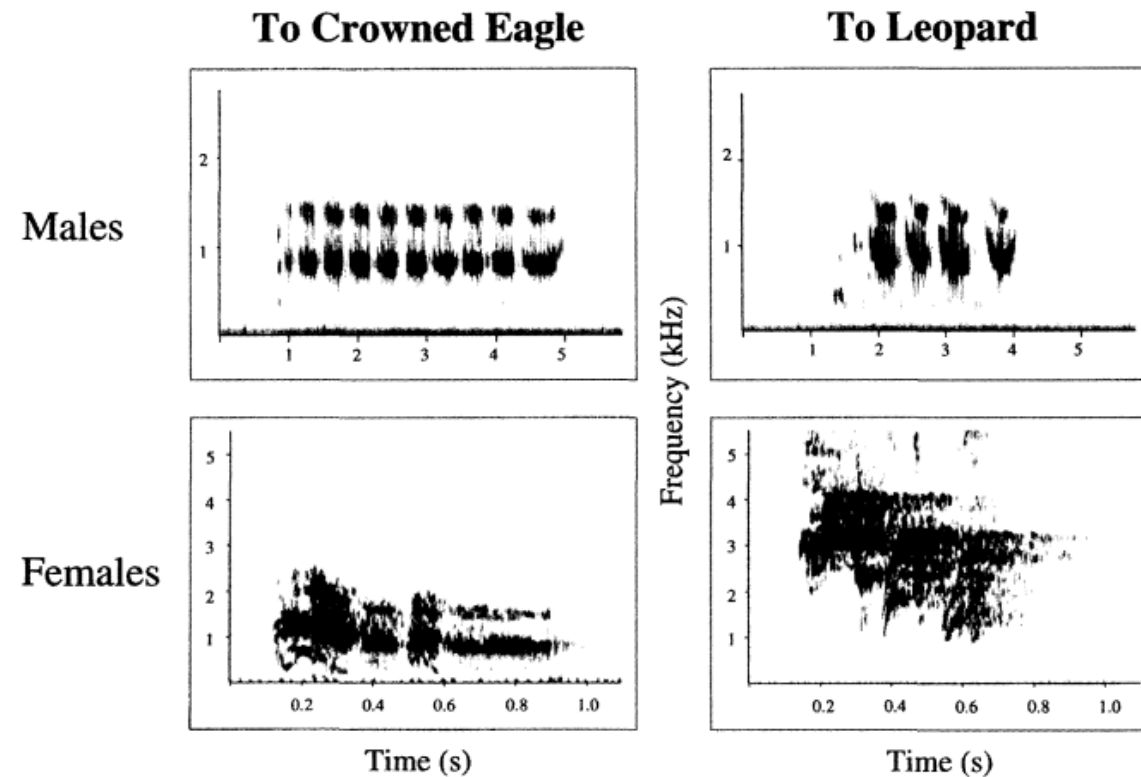


Figure 1. Alarm calls of male and female diana monkeys in response to crowned eagles and leopards.

- Проиллюстрировать эти отличия можно на примере локативных высказываний:
- *Стул у стола.*
- *Стул рядом со столом.*
- *Стул напротив стола.*
- *Стул возле стола.*



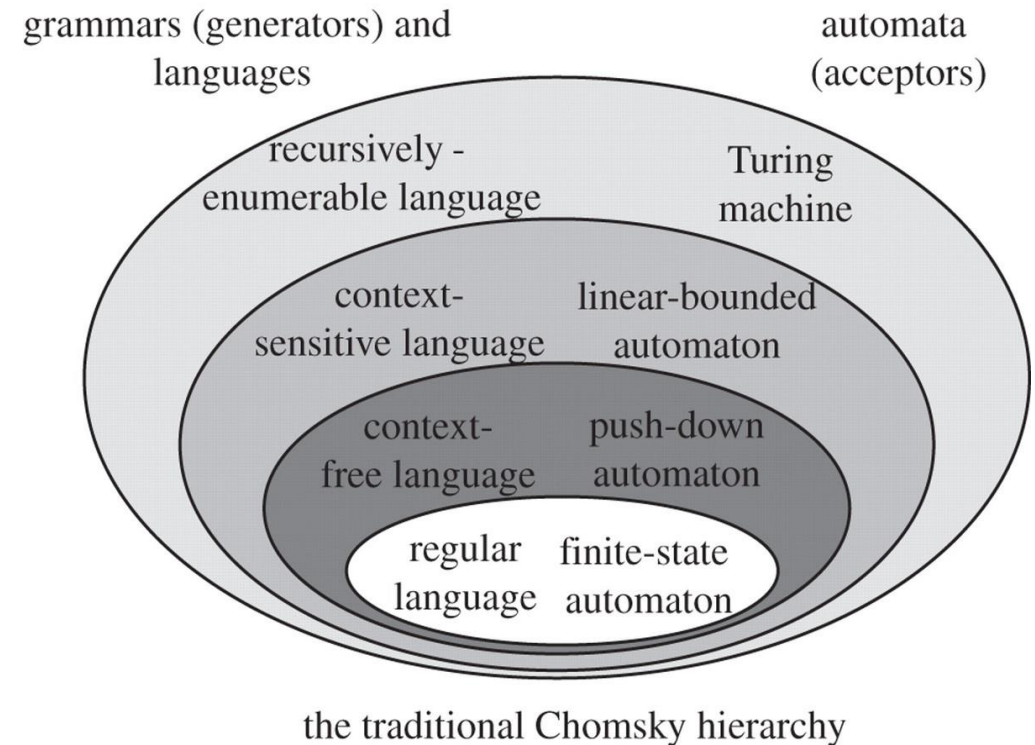
О чем говорят предлоги?

- Предложение «Стул напротив стола» содержит непрямой семантический вывод *«Стол стоит возле стены. У стола есть лицевая сторона, напротив которой стоит стул»*.
Употребление предлогов «у», «рядом» и «возле» обращается к сознательному опыту употребления:
- «у» – если высота стула ниже уровня столешницы.
- «рядом» – если стул массивный и высота его равна или больше стола
- «возле» – если стул не ориентирован на стол и по внешнему виду не составляет с ним пару.

Инженерное решение

- Задача семантической интерпретации предложных связей всегда состоит в том, чтобы указать их корреляты в логической формуле, представляющей смысл рассматриваемого словосочетания.
- Для реализации этой интерпретации используется логический вид представления знаний и концептуальный словарь.
- Далее создаются строгие предметные описания между «синтаксическим хозяином» и «синтаксическим слугой».
- Для машинного вывода достаточно непротиворечивого описания предметной и логической связи между актантами .

- На наш взгляд, наиболее приемлемым ответом на вопрос о нечеткой семантике естественного языка является тезис об уровнях семантической чувствительности и о наличии базовых нечетких универсалий.
- Естественный язык как вычислительная система связан не с правилами вывода, прописанными в архитектуре мозга, а с прагматическими нюансами значения, которые надстраиваются над универсальной категориальной базой.

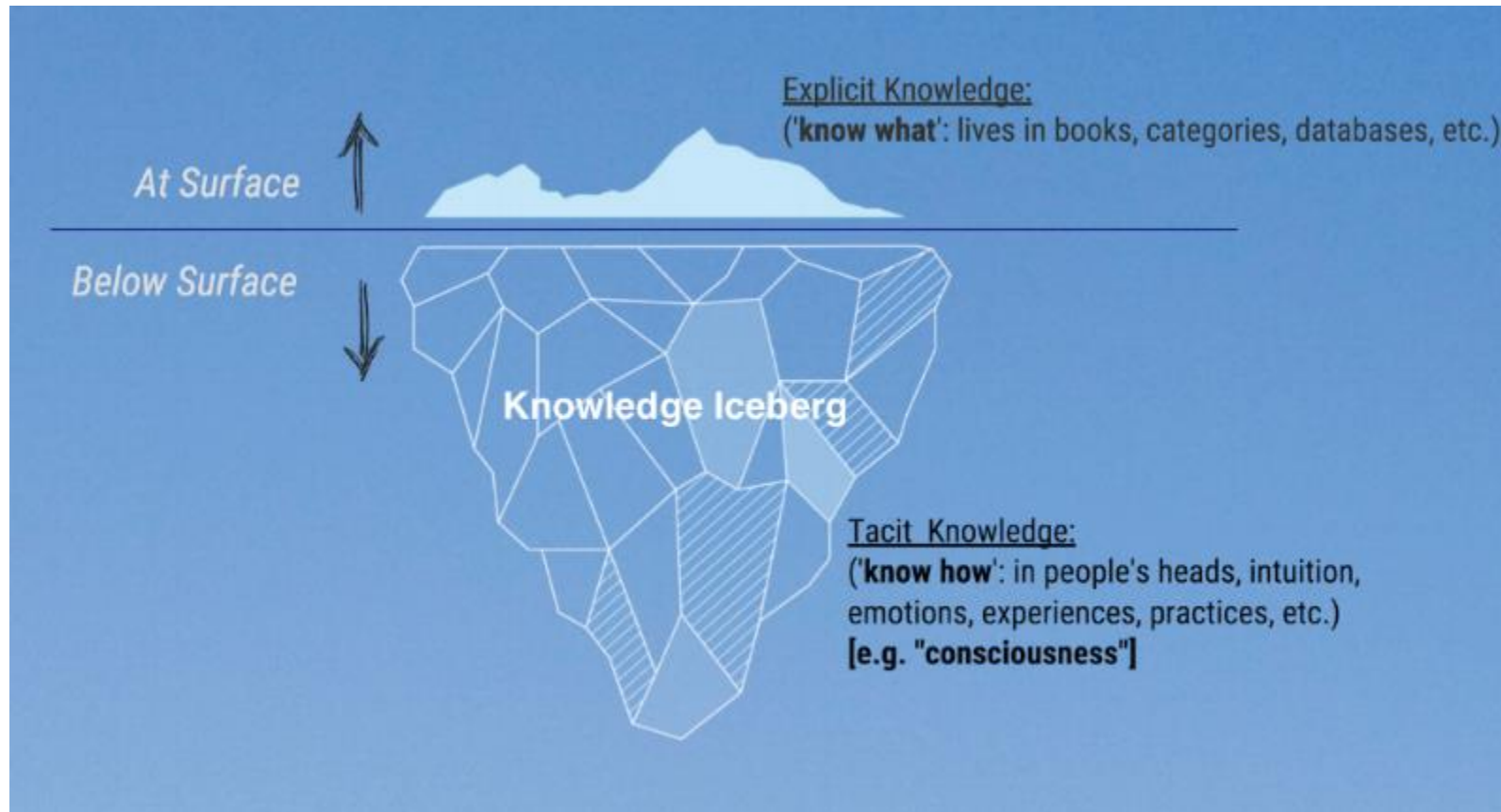


- На основании фреймового анализа некоторые современные авторы предлагают компьютерные функциональные модели сознания [Starjuk J.A., Prasad D.K., 2011]. Всякий раз в подобных работах подчеркивается, что модель призвана **имитировать** биологические процессы обработки сенсорных данных и принятие решений автономными системами.

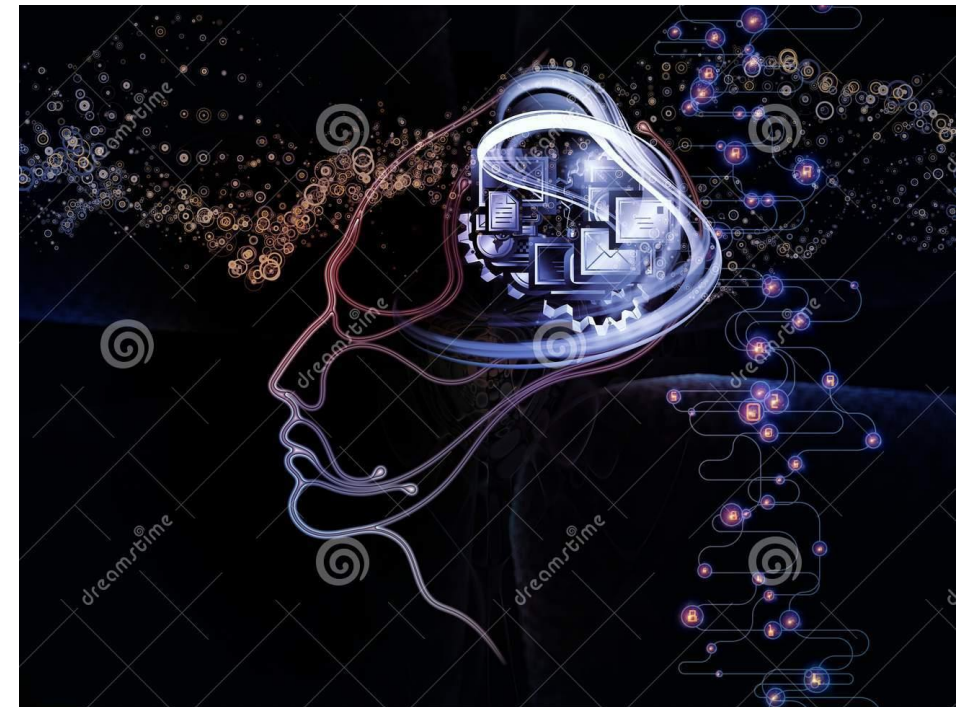
- Подобные модели, как правило, включают в себя сенсорно-моторные блоки, модули эпизодической памяти, самообучающиеся системы и мотивационный процессор.
- Отдельным блоком стоит семантическая память, которая связана с блоком эмоций и субкортикальными вычислениями (Emotion, rewards and sub-cortical processing), с блоком переключения внимания (Attention switching), с сенсорными процессорами (Sensor processors) и с эпизодической памятью (Episodic Memory).

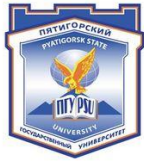
- На наш взгляд, блок семантической памяти – это самое уязвимое место всей вычислительной модели сознания.
- Здесь проходит граница между инженерной и гуманитарной интерпретацией семантических процессов.
- Нечеткость машинных выводов никак не соотносится со спонтанностью и метафоричностью естественной семантики.





- Пока машина работает с данными и информацией, а не со значением и смыслом, противоречивая, но эффективная естественная семантика остается привилегией человека.
- Помимо интуитивной и символической интерпретации, человеку доступна уникальная «база данных» – тексты культуры.
- Эти тексты в свою очередь сопряжены с языковой памятью, сохраненной в языковых привычках, идиомах, фразеологизмах.





Список литературы:

- *Fodor, J.A.* LOT 2. Oxford, New York: Clarendon Press; Oxford University Press, 2008. – 228 p.
- *Olivier, P.; Ananiadou, S.; Maeda, T.; Tsujii, J.* Visualisation: mediating the interchange of information from the verbal to the visual domain. // Mensch und Maschine - Informationelle Schnittstellen der Kommunikation: Proceedings des 3. Internationalen Symposiums für Informationswissenschaft, ISI 1992, Saarbrücken, Germany, November 5-7, 1992. : 1992. – pp. 324–337
- *Starjyk J.A., Prasad D.K.* A computational model of machine consciousness // International Journal of Machine Consciousness, 2011. № 2, vol. 3. – pp. 255–281.
- *Лакофф, Д.* Женщины, огонь и опасные вещи. М.: Гнозис, 2011. – 512 с.
- *Пинкер, С.* Субстанция мышления. Язык как окно в человеческую природу. М.: УРСС; Либрком, 2016. – 560 с.
- *Рубашкин, В.Ш.* Словарная поддержка процедур семантических предложных связей, 2005. <http://www.dialog-21.ru/media/2419/rubashkinv.pdf>.
- *Черняк, Л.* Интеграция данных: синтаксис и семантика. // Открытые системы, 2009. <http://www.osp.ru/os/2009/10/11170978/>.
- *Черняк, Л.* Что делать с хаосом данных? // Открытые системы, 2013. <http://www.osp.ru/os/2013/09/13038279/>.
- *Seyfarth, R.M., Cheney, D.L.* Behavioral mechanisms underlying vocal communication in nonhuman primates. *Animal Learning & Behavior* **25**, 249–267 (1997). <https://doi.org/10.3758/BF03199083>



Благодарю за внимание!

pnbaryshnikov@pglu.ru