

А. А. Глов (компания «Совзонд»)

В 2009 г. окончил Воронежский государственный университет по специальности «география». В настоящее время — руководитель отдела разработки ГИС компании «Совзонд». Кандидат географических наук.

Интеллектуализация геоинформационных систем: подходы и направления

ТЕХНОЛОГИЧЕСКИЕ БАРЬЕРЫ РАЗВИТИЯ ГЕОИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ КАК ПРЕДПОСЫЛКИ ИНТЕЛЛЕКТУАЛИЗАЦИИ

Современные тенденции развития геоинформационных технологий характеризуются наличием ряда технологических барьеров на различных стадиях обработки информации, что ставит вопрос о поиске новых подходов к построению ГИС. В данной работе предпринята попытка произвести анализ основных подходов к построению интеллектуальных геоинформационных систем с позиции теории управления.

Если рассмотреть процесс анализа данных с методологической позиции полной функции управления [5], то можно ввести понятие *полной функции анализа данных*, комплексно описывающей всю цепочку преобразования информации: от получения исходных данных до поддержки принятия решений (рис. 1).

Понятие полной функции анализа данных основывается на том, что конечным результатом любого аналитического процесса является *принятие управленческого решения*. При этом для принятия решения может быть достаточно как поверхностного экспертного анализа исходных данных, так и создания сложного

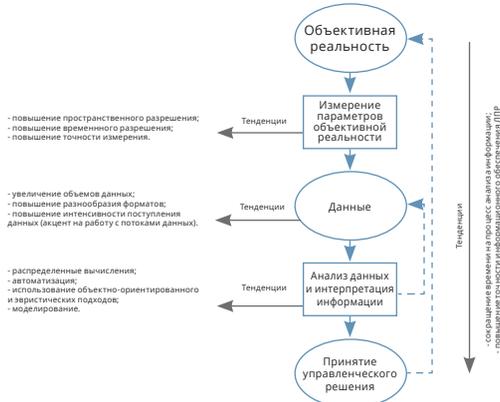


Рис. 1. Структурная схема полной функции анализа данных и основные тенденции развития ГИС на ее стадиях

программно-аппаратного комплекса, ориентированного на информационную поддержку принятия решений и моделирование различных сценариев развития объекта управления.

На наш взгляд, существует две стратегические цели развития геоинформационных и информационно-аналитических систем:

- минимизация времени между измерением параметров объективной реальности и принятием управленческого решения;
- повышение точности и качества аналитического обеспечения принимаемых управленческих решений.

Именно эти две цели определяют основной вектор развития информационно-аналитических технологий. Если рассмотреть тенденции развития информационных систем с позиции полной функции анализа данных на каждом из этапов, то можно сделать следующие выводы:

1. Наблюдается стирание граней между такими традиционными классами информационных систем, как ГИС, системы бизнес-аналитики (Business Intelligence, BI), и различными типами информационно-аналитических систем (ИАС). Данный факт подтверждается, с одной стороны, поддержкой пространственного анализа в системах бизнес-аналитики, с другой — появлением технологий типа Spatial OLAP (пространственный многомерный анализ — технология, характерная для BI-систем). По сути, на рынке более четко выявляется потребность в комплексных информационно-аналитических системах, позволяющих из больших массивов любых разнородных исходных данных формировать аналитический продукт.

2. Измерения параметров объективной реальности характеризуется следующими тенденциями:

- повышение пространственного разрешения измерительного оборудования (применительно к геоданным);
- повышение временного разрешения измерительного оборудования;
- повышение точности и детальности регистрируемых значений.

3. При работе с данными как исходной базой аналитического процесса также отмечается ряд тенденций:

- возрастание объемов данных (тенденция Big Data);
- противоречие между увеличением количества форматов и стандартов представления данных и тенденцией к унификации и абстрагированию форматов данных;
- потоковый характер данных.

4. Этап анализа и интерпретации данных:

- активное развитие технологий распределенных вычислений как следствие проблемы «больших данных»;
- автоматизация стандартных процессов анализа данных;
- возрастание роли объектноориентированного подхода на этапах анализа и интерпретации информации;
- возрастание роли моделирования и эвристических подходов.

Развитие информационно-аналитических технологий в любой момент времени сталкивается с рядом ограничений — так называемых технологических барьеров, обусловленных состоянием развития науки.

В соответствии с этапами полной функции анализа данных это:

- объективные ограничения измерительного оборудования;
- ограничения математического аппарата обработки и анализа информации;
- сложность интерпретации аналитических материалов.

Преодоление вышеуказанных барьеров и пределов в развитии информационно-аналитических технологий возможно за счет использования методов искусственного интеллекта, позволяющего решать сложные неструктурированные задачи.

ПОНЯТИЕ ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ И ИНТЕЛЛЕКТУАЛЬНЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

Существуют различные определения понятия «искусственный интеллект». Так, S. Russel и P. Norvig приводят классификационную таблицу определений данного понятия (табл. 1) [17], включающую четыре группы подходов, описывающих предельные варианты интеллектуализации. В большинстве реальных случаев внедрения технологий искусственного интеллекта

Systems that think like humans (когнитивное моделирование) — когнитивная наука (cognitive science)	Systems that think rationally (законы мышления) — логика
Systems that act like humans (тест А. Тьюринга): — обработка естественных языков (NLP); — инженерия знаний (knowledge engineering); — автоматические рассуждения (automated reasoning); — компьютерное зрение (computer vision); — роботика (robotics)	Systems that act rationally (рациональные агенты) — агентно-ориентированные системы (agent-based systems)

Табл. 1. Классификация определений понятия «искусственный интеллект» [17]

используется гибридный подход, включающий элементы всех четырех групп.

В общем случае понимание искусственного интеллекта как научного направления развивается с двух концептуальных позиций — нисходящей (моделирование когнитивных аспектов деятельности — машинное обучение, инженерия знаний и др.) и восходящей (построение структур данных и алгоритмов, моделирующих биологические системы — нейронные сети, генетические алгоритмы, эволюционные вычисления и др.).

Применительно к информационным системам следует привести еще некоторые

из определений понятия «искусственный интеллект»:

1) «свойство интеллектуальных систем выполнять творческие функции, которые традиционно считаются прерогативой человека» [1];

2) «интеллектуальной называется система, способная целеустремленно, в зависимости от состояния информационных входов, изменять не только параметры функционирования, но и сам способ своего поведения...» [10].

Общий обзор основных направлений искусственного интеллекта приведен в работах [2, 6] (табл. 2).

автоматическое доказательство теорем	theorem proving
автоматические рассуждения	automated reasoning
машинное обучение	machine learning
интеллектуальный анализ данных	data mining
байесовы сети	bayesian networks
генетические алгоритмы	evolutionary computation and genetic algorithms
дедуктивные модели	deductive models
инженерия знаний/управление знаниями	knowledge engineering/knowledge management
интеллектуальные интерфейсы	intelligent interfaces
интеллектуальный интернет (семантический веб)	semantic web
когнитивное моделирование	cognitive modeling
логики (индуктивные, дискриптивные, нечеткие и др.)	logics (inductive / descriptive / fuzzy etc.)
компьютерное зрение	computer vision
нейронные сети	neural networks
онтологии	ontologies
обработка естественных языков и др.	natural language processing

Табл. 2. Основные направления искусственного интеллекта [2, 6]

В общем случае понимание «искусственного интеллекта» включает использование интеллектуальных видов деятельности (распознавание, обучение, память, рассуждение и др.).

Таким образом, среди основных задач интеллектуальных систем можно выделить:

Распознавание — способность системы классифицировать объект или явление, с которым она сталкивается при решении своих задач. Задача распознавания подразумевает наличие некоторых знаний об объектах, с которыми система может встретиться или встречалась на протяжении периода функционирования. Задача распознавания является одной из базовых задач искусственного интеллекта и используется с целью автоматизированного определения стратегии взаимодействия с объектом или явлением, с которым сталкивается информационная система. Данная задача интеллектуальных систем решается в рамках научного направления «теория распознавания образов».

Обучение (в терминах научного направления «машинное обучение») — систематическое обучение алгоритмов и систем, в результате которого их знания или качество работы возрастают по мере накопления опыта [13]. Задача обучения в широком смысле подразумевает возможность создания новых и уточнения созданных моделей в процессе функционирования системы. Задача обучения решается в рамках направлений машинного обучения.

Инженерия знаний — научное направление, изучающее алгоритмы и методы моделирования человеческих знаний, которые могут быть использованы в процессе обучения системы или в процессе логического вывода.

Следует отметить, что решение вышеперечисленных задач может быть использовано на любой стадии аналитического процесса.

ОСНОВНЫЕ НАПРАВЛЕНИЯ И ЗАДАЧИ ИНТЕЛЛЕКТУАЛИЗАЦИИ ГИС

Вопросы интеллектуализации геоинформационных систем привлекают исследователей уже на протяжении достаточно продолжительного интервала времени и рассмотрены в ряде отечественных и зарубежных работ [3, 6, 7, 8, 9, 12, 15, 16, 18].

В качестве базового можно привести определение В. В. Поповича: «Под интеллектуальной принято понимать ГИС, которая включает инструменты и/или подсистемы искусственного интеллекта» [16].

Интеллектуализация ГИС подразумевает использование методов искусственного интеллекта на каком-либо одном либо на всех этапах полной функции анализа данных, что позволяет выделить несколько направлений:

Интеллектуализация процессов поиска и сбора данных. Включает алгоритмы поиска и распознавания данных, с которыми сталкивается система. Данная группа задач включает:

- наличие базы знаний по ресурсам информации;
- методы интеллектуального поиска информации в распределенных системах (семантический веб и др.);
- решение задачи распознавания типов (моделей) данных.

Наличие базы знаний по ресурсам пространственных данных позволяет собирать дополнительную информацию для моделирования и анализа, повышать возможность использования эвристических подходов на основе автоматизированного сбора информации через внешние ресурсы (FTP, HTTP, API и т. д.).

Стратегическим направлением развития базы знаний по ресурсам является использование технологий анализа семантических сетей.

Определение стратегии дальнейшей обработки данных основывается на распознавании типа (модели) данных. Необходимость решения данной задачи связана с тем, что существуют алгоритмы анализа данных как более широкого спектра использования (растровая или векторная модель данных), так и требующие более детального определения модели (тип сенсора для космического аппарата, уровень обработки).

Интеллектуализация процессов анализа и интерпретации данных подразумевает использование эвристических алгоритмов управления данными и алгоритмами анализа, который на основании идентификаций модели данных позволяет:

- в автоматизированном режиме определить стратегию и сценарии анализа (максимальная глубина извлечения информации из источника данных) в форме цепочек и деревьев алгоритмов с учетом логической последовательности и приоритетов алгоритмов;
- подбирать наиболее оптимальные для тематических задач цепочки алгоритмов анализа данных;
- решать задачу распознавания структур или объектов в данных;
- создавать интеллектуальную модель объекта управления;
- осуществлять высокоуровневую объектноориентированную интерпретацию выявленных объектов или явлений;
- выполнять прогнозное и сценарное моделирование развития объекта управления.

Интеллектуализация процессов представления информации. Является достаточно новым и еще слабо распространенным направлением и включает учет и моделирование когнитивных аспектов восприятия информации человеком. Реализация данного направления осуществляется посредством интеллектуальных интерфейсов.

Интеллектуальная ГИС, в отличие от традиционных подходов, характеризуется наличием интеллектуальной подсистемы, центральными звеньями которой выступают база знаний и машина логического вывода.

Использование методов инженерии знаний и построения баз знаний в ГИС основано на использовании трехуровневой абстракции информации: «данные — информация — знания» [14] (рис. 2).

Данные — результаты исходных измерений объективной реальности (исходные факты), представляемые в машиночитаемом виде. Применительно к геоинформационным системам этому понятию соответствуют исходные данные ДЗЗ, цифровые модели рельефа и др.

Информация — обработанные исходные данные, несущие отчетливую информационную нагрузку (данные, переведенные в значимую для пользователя форму). Примером может служить рассчитанный растр индекса NDVI или поверхности уклона рельефа.

Знания — структуры информации/данных, описывающие объекты и процессы объективной реальности, или информация, позволяющая принять управленческое решение [14]. Примером выявления знания может выступать построение модели развития суффозионных процессов в пределах агроландшафтов (рис. 3). Построение подобной модели позволяет анализировать различные сценарии влияния негативных факторов на земельные ресурсы и принять оптимальное управленческое решение.

Следует отметить, что подавляющее большинство современных ГИС пока работает на уровне абстракций «данные–информация», построение основанных на знаниях систем носит, за редким исключением, экспериментальный характер. Наблюдается достижение определенных пределов в развитии современных ГИС. При этом отмечается возрастающая потребность в создании



Рис. 2. Соотношение понятий «данные», «информация» и «знание» в геоинформатике (с использованием материалов [13])

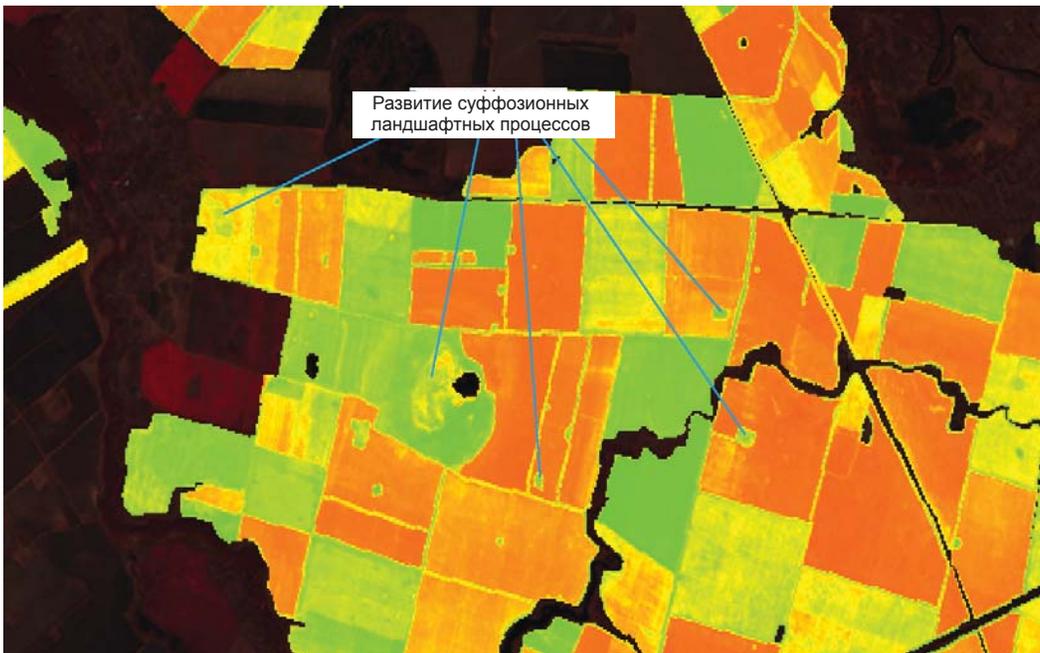


Рис. 3. Пример выявленного знания — развитие суффозионных ландшафтных процессов (Среднехоперский южнолесостепной физико-географический район), позволяет оценить сценарии развития и выполнить планирование мелиоративных мероприятий

интеллектуальных информационно-аналитических систем, позволяющих решать задачи информационного обеспечения принятия управленческих решений на качественно ином уровне. Кроме того, построение интеллектуальных ГИС остро ставит задачу их методологического обеспечения. По нашему мнению, в качестве методологической основы интеллектуализации геоинформационных систем, кроме собственно искусственно-го интеллекта, следует рассматривать такие научные направления как общая теория геосистем [11] и геоинформационное моделирование ландшафта [4].

СПИСОК ЛИТЕРАТУРЫ

1. Аверкин А. Н. Толковый словарь по искусственному интеллекту / А. Н. Аверкин, М. Г. Гаазе-Рапопорт, Д. А. Поспелов. - Толковый словарь по искусственному интеллекту. — М.: Радио и связь, 1992. — 256 с.
2. Гаврилова Т. А. Интеллектуальные технологии в менеджменте: инструменты и системы: учебное пособие 2-е издание / Т. А. Гаврилова, Д. И. Муромцев. - СПб: "Высшая школа менеджмента", 2008. — 488 с.
3. Геоинформатика: Учеб. для студ. вузов / Е. Г. Капралов, А. В. Кошкарёв, В. С. Тикунов и др.; Под ред. В. С. Тикунова. — М.: Издательский центр «Академия», 2005. — 480 с.
4. Глотов А. А. Геоинформационное моделирование эволюции долинно-речных ландшафтов Воронежской области: автореферат диссертации кандидата географических наук / А. А. Глотов. — Воронеж, 2013. — 24 с.
5. Достаточно общая теория управления: постановочные материалы учебного курса факультета прикладной математики-процесов управления СПбГУ. — СПб, 2003. — 419 с.
6. Ивакин Я. А. Интеллектуализация ГИС. Методы на основе онтологий / Я. А. Ивакин. - LAP Lambert Academic Publishing, 2010. — 322 с.
7. Литвинов А. Г. Технология разработки интеллектуальных геоинформационных систем горнопромышленных комплексов / А. Г. Литвинов // Диссертация кандидата технических наук. — М., 2006 г. — 152 с.
8. Попович В. В. Интеллектуальная ГИС в системах мониторинга / В. В. Попович, С. Н. Потапычев, А. В. Панькин, С. С. Шайда, М. Н. Воронин // Труды СПИИРАН. — СПб., 2006. — Т.1, № 3. — С. 172-184.
9. Савиных В. П. Развитие методов искусственного интеллекта в геоинформатике / В. П. Савиных, В. Я. Цветков // Транспорт Российской Федерации, 2010. — № 5. — С. 41-43.
10. Смолин Д. В. Введение в искусственный интеллект: конспект лекций / Д. В. Смолин. — М.: ФИЗМАТЛИТ, 2004. — 208 с.
11. Сочава В. Б. Введение в учение о геосистемах / В. Б. Сочава. — Новосибирск : Наука, 1978. — 320 с.
12. Тикунов В. С. Исследования по искусственному интеллекту и экспертные системы в географии / / Вестн. Моск. ун-та. — Сер. геогр. — 1989. — No 6. — С. 3-9.
13. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / пер. с англ. А. А. Слинкина. — М.: ДМК Пресс, 2015. — 400 с.
14. Kendal S. L An introduction to knowledge engineering / S. L Kendal, M. Green. — London: Springer, 2007. — 287 p.
15. McKeown David M. The role of artificial intelligence in the integration of remotely sensed data with geographic information systems/ David M. McKeown. — Pittsburgh, 1986. — 36 p.
16. Popovich V. Intelligent GIS Conceptualization / Vasily V. Popovich // Information Fusion and Geographic Information Systems, Lecture Notes in Geoinformation and Cartography, 2014. — P. 17-44.
17. Russel Stuart J. Artificial Intelligence. A modern approach / Stuart J. Russel, Peter Norvig. — New Jersey, 1995. — 932 p.
18. Voženilek V. Artificial intelligence and GIS: mutual meeting and passing / V. Voženilek // 2009 International Conference on Intelligent Networking and Collaborative Systems, 2009. — P. 279-284.