

Веб-майнинг (Web Mining)

Разделы: [Бизнес-задачи](#)

Веб-майнинг — направление интеллектуального анализа данных, в рамках которого решаются задачи автоматического обнаружения веб-документов и услуг во Всемирной паутине, извлечения информации из веб-ресурсов и выявления зависимостей и закономерностей в массивах данных из Интернета.

Другими словами, технология Web Mining применяет технологию Data Mining для анализа неструктурированной, неоднородной, распределенной и значительной по объему информации, содержащейся на Web-узлах.

Системы Web Mining могут ответить на многие вопросы, например, кто из посетителей сайта является потенциальным клиентом интернет-магазина, какая группа клиентов интернет-магазина приносит наибольший доход, каковы интересы определенного посетителя или группы посетителей.

В Web Mining можно выделить следующие этапы:

- входной — получение «сырых» данных из источников (логи серверов, тексты электронных документов);
- предобработки — преобразование данных к форме, необходимой для успешного построения той или иной модели;
- моделирования — построение и применение аналитических моделей;
- анализа — интерпретация полученных результатов.

Конкретные процедуры каждого этапа зависят от поставленной задачи. В связи с этим выделяют различные категории Web Mining:

- Анализ использования веб-ресурсов (Web Usage Mining).
- Извлечение веб-структур (Web Structure Mining).
- Извлечение веб-контента (Web Content Mining).

Общая взаимосвязь между категориями Web Mining и задачами Data Mining изображена на рисунке:



Примерами задач бизнес-аналитики, в которых могут быть применены технологии Web Mining, являются:

- описание посетителей сайта (применяются методы кластеризации, классификации);
- описание посетителей, которые совершают покупки в интернет-магазине;
- определение типичных сессий и навигационных путей пользователей сайта (применяются ассоциативные правила, последовательные шаблоны);
- сегментация рынка интернет-услуг.