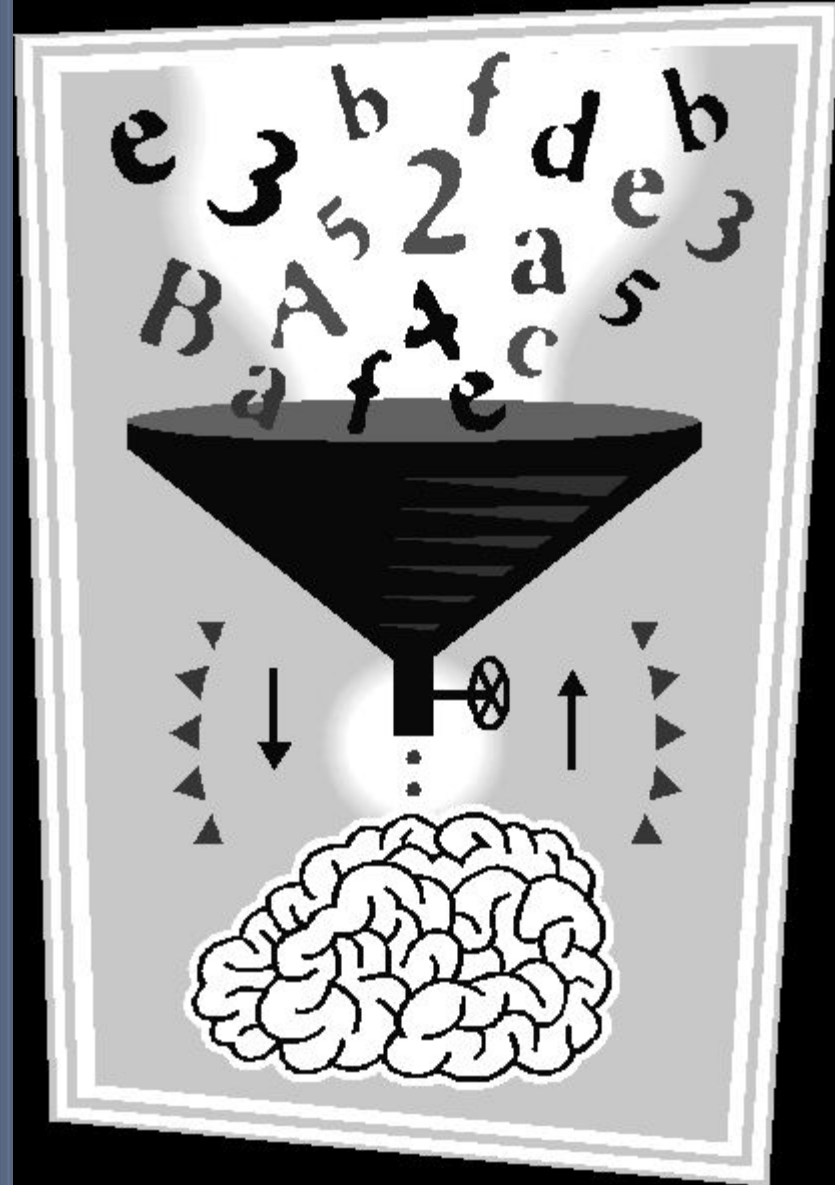


# ПОТЕР Я ЗНАНИ Й.




# Жизнь в мусорной куче, или нужны ли нам знания?

Управление знаниями и интеллектуальный капитал — это громкие слова, породившие тысячи книг, статей и конференций... Но еще никто из практически работающих в этой области не смог оценить, увеличиваются ли активы знания или уменьшаются. Эти оценки показывают только то, что управление знаниями эволюционирует от волшебства к астрологии...


*Пол Страссман*





Очерки о базах знаний и интеллектуальной обработке вызывают закономерный скепсис — в них мало объективности и технически точной информации, зато много двусмысленности, а иногда и безграмотности. Названия «интеллектуальный», «система обработки знаний» ставятся на множество изделий. Но на вопрос «а почему, собственно?» никто из производителей не может внятно ответить.

Сам Страссман не определяет (в пределах статьи) знания и интеллектуальность, но дает неплохую бизнес-метрику для их денежной оценки. Однако его высказывания подразумевают знания как человеческий актив и умалчивают о знаниях как предмете технологической обработки. Из его слов очевиден лишь факт, что и носителем интеллектуальных знаний, и инструментом их обработки является человек.



# Добро пожаловать в мир хаоса!



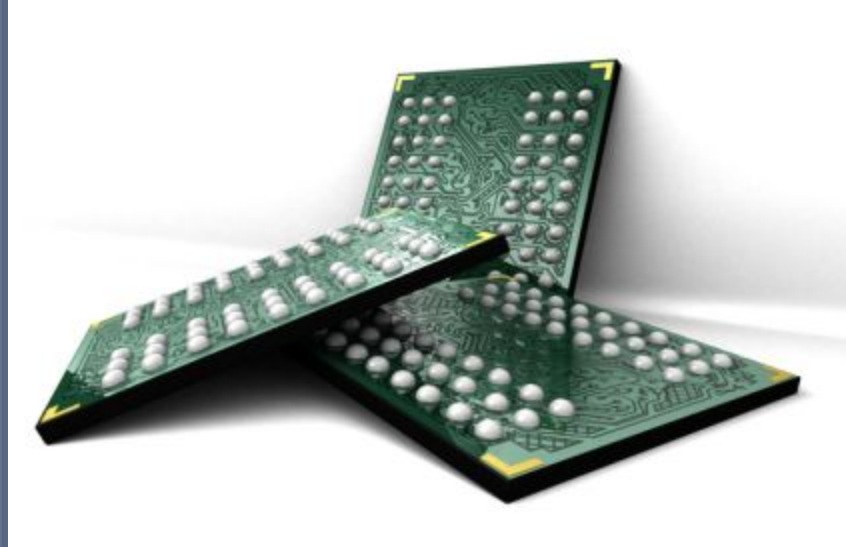
Информационный хаос прогрессирует — лавинообразный рост объемов информации и ее обесценивание сопровождаются соразмерным падением объема знаний. Этот процесс имеет две веские причины.

Первая является техногенной и заключается в резком удешевлении доступа к данным и ее носителям, в том числе и к Internet. Это позволило тиражировать информацию в огромных, ничем не ограниченных количествах. Именно благодаря Сети бизнес стал массово получать деньги за информацию, а не за знания. Выходя на просторы Сети, этой великой мусорной кучи, чувствуешь себя копошащимся компьютерным бомжом — а вдруг найдешь то, что нужно?



Вторая причина заключается в резком падении качества и ценности технической информации. Истоки этого в том, что за последние лет десять изделия ИТ (включая электронную информацию) как класс товара, стали все более приближаться к ширпотребу. Соответственно изменились и методы сбыта ИТ-изделий, которые теперь включают все рекламные трюки психологической «ломки» покупателя, основанные на знании его вкусов, предпочтений и слабостей. К тому же изделия ИТ, особенно информация, — бурно изменяющийся класс товара с коротким жизненным циклом. Не успеешь продать сегодня — не успеешь никогда; неудивительно, что принципы «и так сойдет» и «пока живут на свете дураки...» процветают в ИТ-индустрии, как нигде. Многие производители публично возводят их в ранг своей бизнес-политики и весьма успешно практикуют с согласия большинства ИТ-потребителей, предпочитающих плохое хорошему. Применительно к информации знания подменяются схоластикой. Объективно этому способствует дефицит времени как производителя (на качество производства ИТ-изделий), так и покупателя (на качество потребления).

Чтобы понять происходящее в ИТ, важно выяснить, чем же знания принципиально отличаются от информации и данных.



# Что это такое?

Начнем с попытки дать определения терминам «знание», «информация», «данные». Сложность этой задачи заключается в следующем.

Общепотребительность этих терминов и их кажущаяся синонимичность. Отличаются друг от друга определения этих терминов в различных справочниках, в том числе и прошедших серьезное научное рецензирование. Большинство этих определений описывают понятия предметных областей (например, «документальная информация», «коммерческая информация») и мало дают представления о собственных свойствах знаний, информации и данных как таковых.

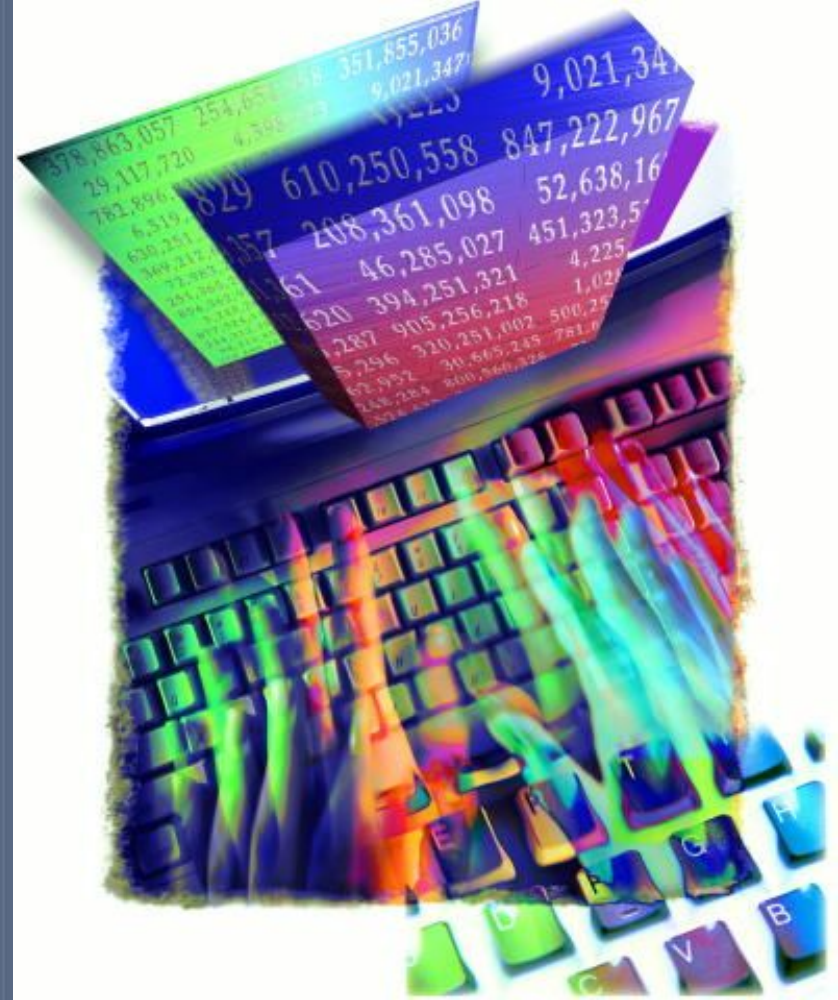
Взаимная трансформация терминов «знания-информация», «информация-данные». Этому явлению уделим особое внимание, так как граница между ними даже для большинства ИТ-специалистов весьма условна.

Миграция понятий этих терминов от общепотребительного смысла к информационно-технологическому. Обусловлена возросшим влиянием ИТ. Если раньше знания традиционно являлись лишь средством описания предметных областей, то сейчас сами стали предметом изучения и обработки («знания о знаниях»).

Теперь — определения (за основу была взята часть справочных определений, переработанных автором в целях устранения неточностей).

# Данные.

Данные — сведения, представленные в определенной знаковой системе и на определенном носителе для обеспечения возможностей хранения, передачи, приема и обработки этих сведений. Уходя от способов представления, хранения и передачи, данные можно рассматривать как абстрактную субстанцию, несущую некоторую информацию. Данные безотносительны к содержанию информации.



# Знание.

Знание — проверенные общественной практикой полезные сведения, которые могут многократно использоваться людьми для решения тех или иных задач. Обратите внимание на «рыхлость» этого традиционного определения в сравнении с нижеследующим.

Знание как объект коммерции и автоматизации — логически полный ограниченный набор сведений для непосредственного решения требуемой задачи (ряда задач) подготовленными специалистами. Такие сведения выражаются в системе понятий, принятой в рамках некоторой науки или производственной деятельности, и имеют стандартное представление. Ограниченный набор позволяет задать уровень подготовки специалистов (обладание требуемым уровнем понимания).





# Информация.

- ❖ Полезная информация — набор сведений (разъяснений, сигналов), уменьшающих степень неопределенности у их получателя. В отличие от знаний этот набор может быть логически неполным.
- ❖ Релевантная информация (недостающие знания). Полезная информация, полностью устраняющая степень неопределенности у ее получателя.
- ❖ Информационный мусор — данные, не несущие полезной информации и многократно увеличивающие временные и прочие издержки пользователя на извлечение и обработку полезной информации.
- ❖ Информация (в общем случае) — совокупность знаний с информационным мусором.



Обратите внимание на степень неопределенности. Она является той чертой, которая мешает провести границу между релевантной информацией и знаниями согласно их определениям. Она, как антипод уровню подготовленности, является индивидуальным свойством человека и его состояния, и именно она влияет на взаимную трансформацию терминов «знания—информация—мусор». Если после получения человеком сведений степень неопределенности становится равна нулю, то информация в этом сообщении является стопроцентной релевантной информацией, если же  $S.H. = 1$ , то информация является стопроцентным мусором.



# Реальность: обезценивание информации и потеря знаний.

Сегодня по интересующему вопросу вы можете обнаружить множество источников сведений и получить множество порций разнородной информации, вес знаний в которых крайне мал. Но даже лишь у мизерной части этих знаний семантика формализована и обеспечена стандартным доступом в виде известных форматов и интерфейсов.



# Причины.

1. Легкая генерация информации. На одном наборе знаний можно изготовить множество наборов информации. Степень неопределенности, допускаемая при чтении информационного сообщения, позволяет описывать информацию множеством способов. Если за информацию платят, то почему бы этим не воспользоваться? Например: берем справочные сведения, перефразируем, разбавляем приятными читателю эмоциональными фразами и — получите очередную книжку!



# Причины.

2. Легкое засорение знаний. Нет проблем с засорением знаний и их постепенным превращением в информацию. Понимание, как семантический эффект, возникает главным образом при чтении текстов. Тексты выражены средствами естественных языков, которые легко объединяют и техническую информацию, и эмоциональное состояние человека. Этим объединением стала страдать современная техническая литература, сильно разбавляемая эмоциями авторов книг. При этом удельный вес технической части в книгах и ее качество падает. Печально, но факт: мы платим за бумагу, ее объем (и книжные полки под эту макулатуру), за избыток потраченного времени на ее чтение и обработку, и лишь в мизерной части — за полученные знания. Все это в полной мере относится и к информации в электронном виде.

# Причины.

3. Отсутствие средств единого синтаксического и семантического представления знаний. Посмотрите хотя бы на ИТ-специалистов. Они — живое воплощение отсутствия таких средств! Сегодня весьма немногие из них понимают сущность знаний и умеют ясно и доступно излагать их, особенно в виде технически грамотных спецификаций. Отсюда множество проблем, достойных отдельной статьи.
4. Сложность семантической и синтаксической верификации знаний. Мизерное изменение текста может легко изменить смысл знаний.
5. Например, перестановка запятой в известной фразе «казнить нельзя, помиловать» при отсутствии или неполноте смыслового контекста у фразы.
6. Легкая потеря семантической полноты. Достаточно потерять запятую в примере «казнить нельзя, помиловать» (например, при распознавании через сканирование), и знания превращаются в мусор.

# Добыча и использование знаний.

После всего вышесказанного уместен вопрос: что же мы собрались добывать? информацию? полезную информацию? знания? Сможем ли мы их грамотно описать и эффективно использовать? Как заставить работать интеллектуальный потенциал работников фирмы, который, как убедительно показал Страссман, сопоставим не только со стоимостью ее материальных активов, но и нередко кратно их превышает?





КОНЕЦ?