

Web Data Extraction and Knowledge Graph Management: Challenges and Opportunities

Dr Andrey Kravchenko

Department of Computer Science, University of Oxford



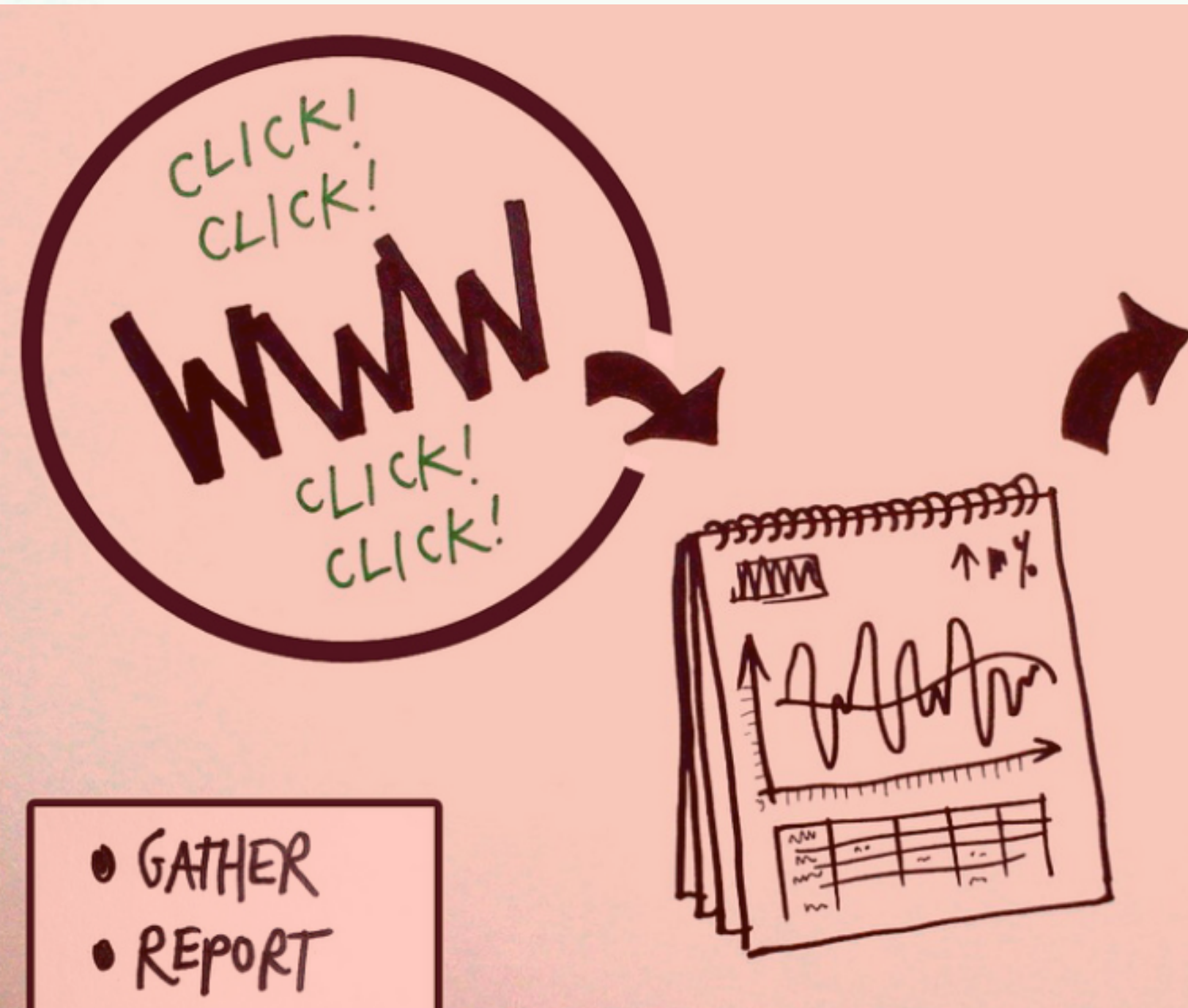
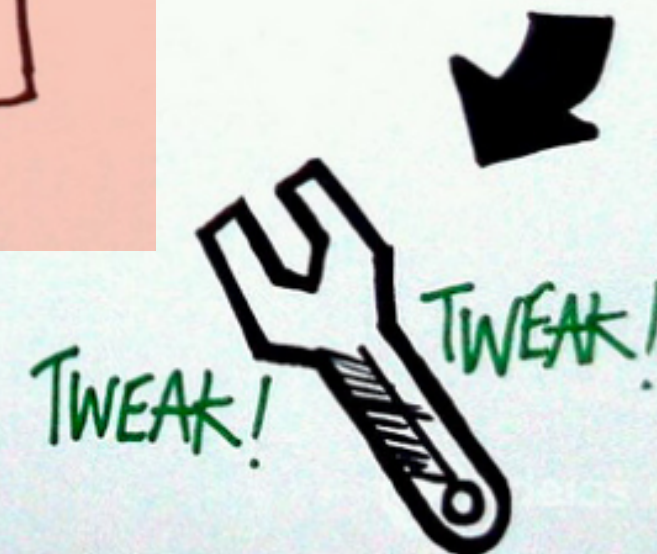
Web Data Extraction and Knowledge Graph Management: Challenges and Opportunities

- Web data extraction has gained prominence in the last two decades, Knowledge Graph (KG) Management in the last decade.
- Both tasks are of major importance for all companies that have to process and structure large amounts of heterogenous data from different data sources, including the web.
- These two tasks are often interconnected.

Overall structure of the talk

- Introduction
- Web Data Extraction: Problem and State-of-the-Art for Approaches.
- DIADEM: Domain-Centric System for Fully Automated Web Data Extraction
- DIADEM demo.
- Knowledge Graph Creation and Management: Problem and State-of-the-Art Approaches.
- VADA: a Novel System for Data Wrangling and KG Creation and Management.
- DeepReason.ai: an Oxford University Spinout, a KG platform with Unparalleled Reasoning Power.

Data Extraction





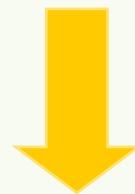
“Without big data analytics, companies are blind and deaf, wandering out onto the Web like deer on a freeway.”

—GEOFFREY MOORE

The McKenna Group, VC, Author of Crossing the Chasm & Inside the Tornado

Data Extraction

The screenshot shows the FindersKeepers website interface. At the top, there's a navigation bar with links: About Us, Contact Us, Careers, My Shortlist, and Login / Register. The main header features the FindersKeepers logo and a badge stating 'Sunday Times 2012 Letting Agent of the Year'. Below the header is a search bar with a 'GO' button. The main content area is divided into two columns. The left column contains search filters: '1. Minimum Bedrooms?' with buttons for 1, 2, 3, 4, and 5+; '2. Select Your Monthly Rent: £550 to £3000 & up' with a slider; '3. Where?' with a map of Oxford and surrounding areas (Banbury, Chipping Norton, Witney, Burford, Abingdon, Didcot, Wallingford, M40, Thame, Bicester, Wheatley) and buttons for 'Map Search' and 'List Search'; and a 'Search Properties by School' section with dropdowns for 'Select School Type' and 'Select School', and a 'Go' button. The right column features a large image of a modern living room with a fireplace and large windows. Overlaid on the image is text: 'Just let... Outstanding apartment in the Brew Tower development, Abingdon.' and a 'Find out more' button. Below the image, there's a section titled 'Sunday Times letting agent of the year' with a brief description of the company and a link 'As featured in'.



ref-code	postcode	bedrooms	bathrooms	available	price
33453	OX2 6AR	3	2	15/10/2013	£1280 pcm
33433	OX4 7DG	2	1	18/04/2013	£995 pcm

Information Extraction

Text, lots of it

Subscriptions to Stock in the New Church.

WHEREAS THE TRUSTEES OF THE FIRST PRESBYTERIAN SOCIETY, in the late town, now city, of Buffalo, have resolved to erect a New Church Edifice on the site of the old building, now belonging to said Society, as soon as an amount shall be subscribed which shall be deemed sufficient by the Trustees to enable them to settle and arrange with the present Pew holders, and to erect, complete and furnish the proposed New Edifice, including an organ.

Now FOR THIS PURPOSE, We, the subscribers, do severally promise to pay to the Treasurer of the First Presbyterian Society, in the late town, now city, of Buffalo, the sums set opposite to our respective names, to be applied towards the building and completion of a New Church Edifice, on the lot now owned by said Society, after the plans furnished by Messrs. Brown & Otis, or such modifications of the same as the Trustees shall deem expedient, and for the payment of such subscriptions we severally agree, whenever One Hundred Thousand Dollars are subscribed hereto, including such amount as present Pew holders may consent to receive in certificates as hereinafter mentioned, to give our promissory notes for the amount of our respective subscriptions, payable at one of the Banks in Buffalo, in six equal payments—at four, eight, twelve, sixteen, twenty and twenty-four months after the date thereof.

The Church when completed, is to be dedicated and set apart exclusively as a place of public worship, according to the rites and forms of the Presbyterian Church of the United States of America; and the subscribers hereto may become interested therein in the manner and upon the terms hereinafter specified; to all of which the corporation above named, by the acceptance of our respective signatures to this subscription, declare their assent, and to bind themselves, and their successors in office, to the just and faithful performance of the same.

When the said New Church Edifice shall be completed, or sooner if the Trustees shall deem it expedient, the Trustees shall affix a valuation to each of the Slips in it, except those reserved as free seats, and shall, upon due notice, offer the Slips, to which valuations are affixed, for competition of choice at public auction, subject to such conditions and reservations as the Trustees may deem proper for the future maintenance of public worship therein, and to defray the contingent and other expenses of said corporation.

At the time of giving the notes above specified, transferable certificates shall be issued by said corporation for the same amount, signed by the President and countersigned by the Clerk; which certificate, when said notes shall be fully paid, shall entitle the holder to apply the amount thereof as cash in payment for Slips in said Church, at the first public sale above specified. And after the first public sale, the holder of any such certificate may select any unsold Slip at the valuation so affixed, and apply the amount of such certificate, when fully paid, as cash in payment therefor, on such terms and conditions as the Trustees shall prescribe; but all such certificates, not presented and exchanged for Slips within six months after the Church shall be dedicated, shall be deemed a donation to the Church.

It is understood and agreed by the parties hereto, that the Trustees are to make efforts to have the present Pew holders release their interest therein respectively at the original appraised value thereof, or such other valuation as shall be agreed upon between them, and to receive therefor transferable certificates substantially in the same form as the certificates above mentioned; and the amount so issued for pews released shall be deemed a part of the above subscription of One Hundred Thousand Dollars; but this subscription shall not be obligatory upon us, unless the said sum of One Hundred Thousand Dollars shall be subscribed as above provided, within one year from the date hereof.

DEAR SIR:—The above is the form of the subscription to the Stock in the New Church, as adopted by the Trustees. You are requested to examine it, and be prepared to act when called upon by the Committee appointed to circulate the Book for Subscriptions. The Committee will have an onerous duty to perform, and it is hoped that you will aid in facilitating their labors as far as practicable.

Very Respectfully,
H. SHUMWAY,
President of Board of Trustees,

THE SILVER CHAIR

by C.S. Lewis

CHAPTER ONE
BEHIND THE GYM

...nn day and Jill Pole was crying behind the gym.

...ause they had been bullying her. This is not going to be a school story, so I ... possible about Jill's school, which is not a pleasant subject. It was "Co-ool for both boys and girls, what used to be called a "mixed" school; some ... ly so mixed as the minds of the people who ran it. These people had the idea ... should be allowed to do what they liked. And unfortunately what ten or ... st boys and girls liked best was bullying the others. All sorts of things, horrid ... ich at an ordinary school would have been found out and stopped in half a ... ool they weren't. Or even if they were, the people who did them were not ... d. The Head said they were interesting psychological cases and sent for ... them for hours. And if you knew the right sort of things to say to the Head, ... that you became rather a favourite than otherwise.

...Pole was crying on that dull autumn day on the damp little path which runs ... f the gym and the shrubbery. And she hadn't nearly finished her cry when a ... e corner of the gym whistling, with his hands in his pockets. He nearly ran ...

...ere you're going?" said Jill Pole.

...e boy, "you needn't start -" and then he noticed her face. "I say, Pole," he ...

...s; the sort you make when you're trying to say something but find that if you ... ying again.

...se - as usual," said the boy grimly, digging his hands farther into his pockets.

...was no need for her to say anything, even if she could have said it. They both ...

...said the boy, "there's no good us all -"

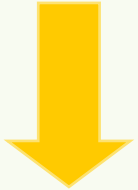
...he did talk rather like someone beginning a lecture. Jill suddenly flew into a ... uite a likely thing to happen if you have been interrupted in a cry).

cca Schneider

'lick, Shudder: The 'Document Performance'.

...erformance theorist and theatre historian, I have been engaged for a ... me in thinking not only about theatricality and its composition in ... tion, but about that vexed category of the 'live' as a descriptor of ... mance—one often used to distinguish theatre or performance art from ... media, such as photography. Recently, I've been engaged in trying to ... ik some of our long-sedimented distinctions between medial forms ... as between performance and photography) as I have been suspect of ... tegory of the 'live', to the degree that it pretends to delimit a present ... nt to instantaneity. I have also been suspect of the habit of the ... rry of the 'live' to reify a supposed antithesis: very often posited as the ... , the remains, the trace, or death. Supposedly, a photograph is always ... y a record, as it appears to survive something that can be called the ... vent—standing in as a trace or document of something that 'was there' ... rthes would have it) but 'is' no longer. Does this way of thinking about ... graphy limit our access to a photograph as event—as a performance of ... on—taking place 'live' in an ongoing scene of circulation, re-circulation, ... nter, re-encounter, and collaborative exchange with viewers, ... rers, reenactors, re-performers, re-photographers?

...It may be that thinking about performance as taking place only in a ... and disappearing present limits our thinking in significant ways. Why ... performance not take place as a photograph? We are habituated, for ... ple, to thinking of the 'present' as singular, unfolding in a linear



subject	predicate	object
Trump	headOfState	USA
Elisabeth II	headOfState	UK

WHY?



Data Extraction feeds

Big Analytics and Verticals



McKinsey&Company


BLACKLOCUS
PRICE TO WIN

M
Man

LAHL *gh* **FRM**
Innovating to perform

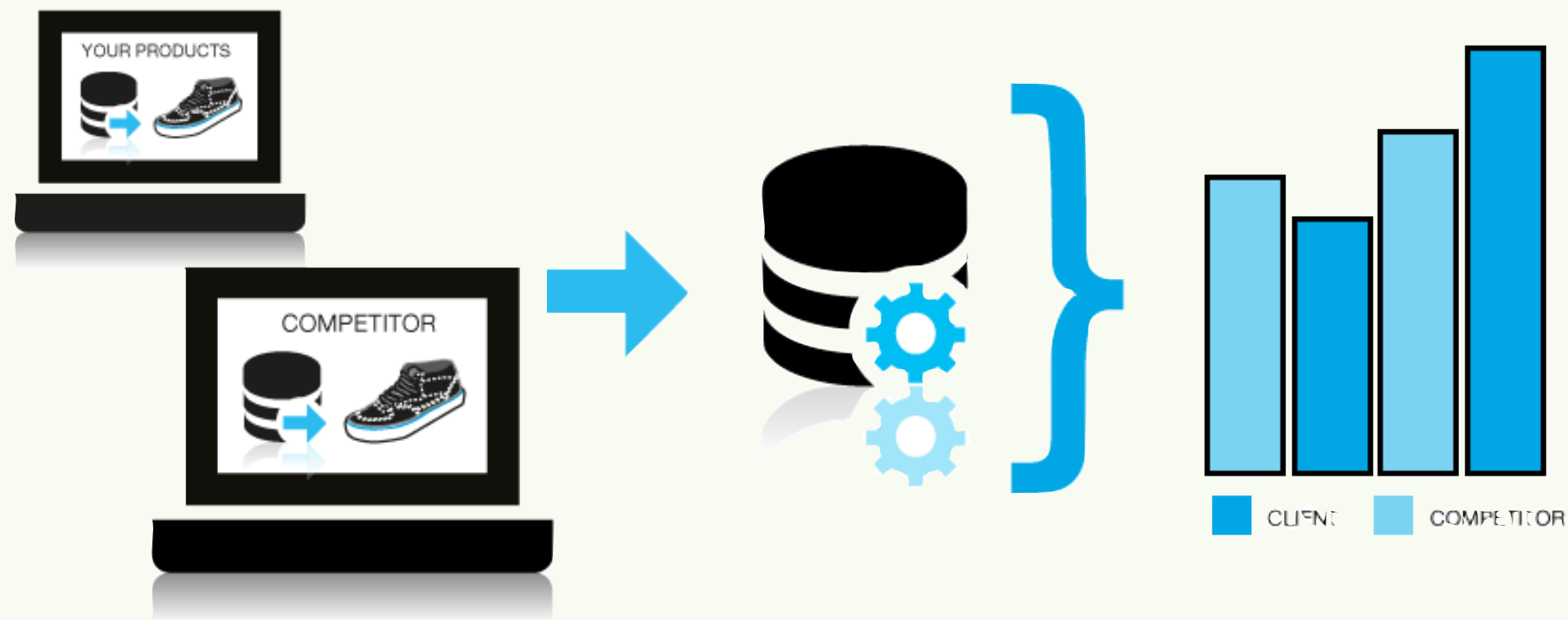


nestoria



H O R U S
S E C U R I T Y
CONSULTANCY LTD

Price Intelligence



Price intelligence :=

—WIKIPEDIA

- ① **Discovery:** finding product pages on competitor sites
- ② **Extraction:** gather price, shipping and availability data
- ③ **Matching:** determining comparable products in your own catalogue

Price Intelligence



“plan to use competitive price intelligence”

–2013 RSR PRICING BENCHMARK

- **Home depot** bought data extraction startup **Blacklocus**
- **McKinsey** bought data extraction company **Lixto**
- **Successful funding** for several data extraction startups

Price guarantees

- detect **violations** of price guarantee
- track contingent trends

hotel.de
check & book

Hotel reservation Meetings Packages For hotels For companies For companies

Homepage | Vicinity search | Cancellation | Reservation by phone | G

Compare hotels

Hotel Name ▲▼

☐ Ethos Hotel
59 Western Road
OX1 4LF Oxford, United Kingdom
6,4 o

☐ Hawkwell House Hotel
Church Way
OX4 4DZ Oxford, United Kingdom
7,6 o

☐ Holiday Inn Express Oxford-Kassam Stadium
Grenoble Road
OX4 4XP Oxford, United Kingdom
7,3 o

☐ Holiday Inn OXFORD
Peartree Roundabout
OX2 8JD Oxford, United Kingdom
7,5 o

☐ Cotswold Lodge Classic Hotel
66a Banbury Road
OX2 6JP Oxford, United Kingdom
7,6 o

Hotel.de
simply book for less

Price Comparison

Hotel: _____

Email: _____

Outbox: _____

Sent Email: _____

Email Template: _____

Tools

Hotel Matching

Storno Mapping

Console

Price Comparison

HMID: _____

Hotel name: _____

City: _____

Nights: _____

Room cat.: _____

Escalation level: _____

Advertisement: _____

Arrival date: _____

Price variance: _____

Country: _____

Best price: ☐

Breakfast: ☐

Price violation: ☐

Matched: ☒

Sorting: _____

Search: _____

Reset: _____

1	2	3	...	155							
Date	HMID	Hotel	ES	Source	Arrival	Nights/PAX/Room	Cancellation	Breakfast	St	Price	Diff.
17.04.2008	102521	artgenossen Landarhotel Linder - DE	1	hotel.de	18.05.2008	1 N. / 2 Adult / Doppelzimmer	0T	U +7,80	1	86,00	
				booking	18.05.2008	1 N. / 2 Adult / Doppelzimmer	0T	U +7,80	1	61,00	-24,6% / -25,00
17.04.2008	144745	Insel Hotel Bonn - DE	1	hotel.de	18.05.2008	1 N. / 2 Adult / Doppelzimmer	NO_MATCH	UF	1	120,00	
				hrs	18.05.2008	1 N. / 2 Adult / Doppelzimmer	1T	UF	1	99,00	-17,5% / -21,00
17.04.2008	55489	Hotel Berlin Sindelfingen - DE	2	hotel.de	18.05.2008	1 N. / 2 Adult / Design Doppelzimmer	0T	UF	1	100,00	
				booking	18.05.2008	1 N. / 2 Adult / Standard Doppelzimmer	0T	UF	1	89,00	-10,3% / -20,00
17.04.2008	214437	Hotel Oase	1	hotel.de	18.05.2008	1 N. / 2 Adult / Doppelzimmer	0T	UF	1	65,00	

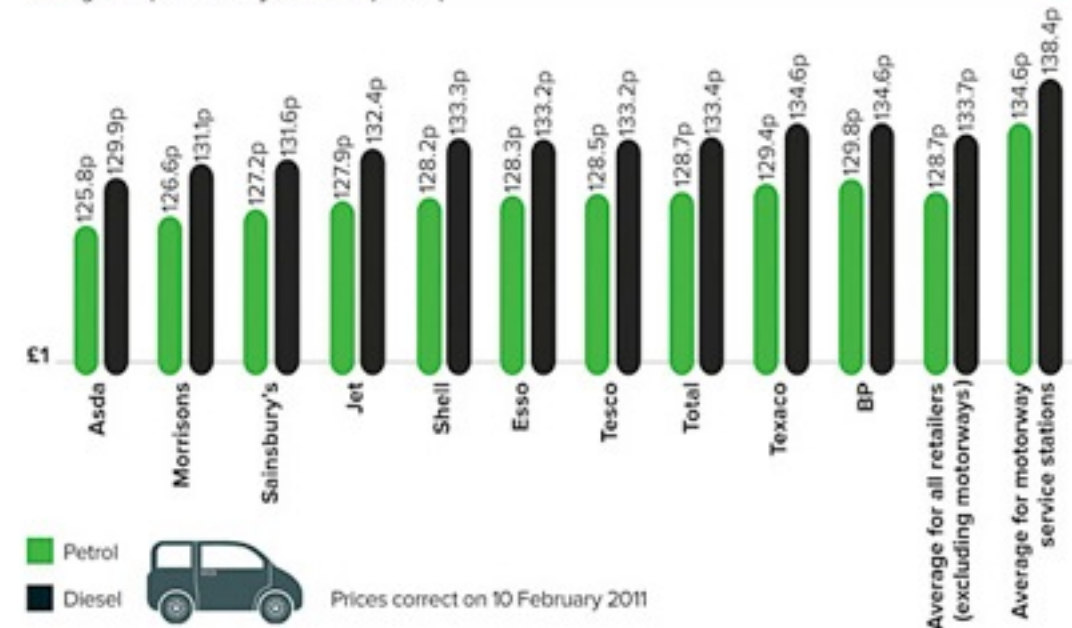
Market intelligence

- online market intelligence
- replace manual or semi-supervised collection



HOW THE BRANDS COMPARE

Average fuel prices at major retailers (ranked)



Financial Analytics

Hypothesis: Big data trumps manual curation

- Result:
- With MAN analysts: **agricultural** futures
 - With 40 web sources: **94%** trend prediction accuracy
 - Beats manually crafted algorithm

House price index

TOP
10
SMALL
TOWNS

FIXR

BEST PLACES TO BUY A HOUSE

FIXR.COM wants to illustrate the top 10 best small cities to live. The ranking takes into account income, population, education, leisure, culture & more.



MEDIAN HOME PRICE



SUNNY DAYS PER YEAR



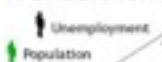
LEISURE & CULTURE



EDUCATION



POPULATION



AVERAGE PROPERTY TAX



MEDIAN FAMILY INCOME



A History of Home Values

The Yale economist Robert J. Shiller created an index of American housing prices going back to 1890. It is based on sale prices of standard existing houses, not new construction, to track the value of housing as an investment over time. It presents housing values in consistent terms over 116 years, factoring out the effects of inflation.

The 1890 benchmark is 100 on the chart. If a standard house sold in 1890 for \$100,000 (inflation-adjusted to today's dollars), an equivalent standard house would have sold for \$66,000 in 1929 (66 on the index scale) and \$199,000 in 2006 (199 on the index scale, or 99 percent higher than 1890).

DECLINE AND RUN-UP: Prices dropped as mass production techniques appeared early in the 20th century. Prices spiked with post-war housing demand.

BOOM TIMES: Two gains in recent decades were followed by returns to levels consistent since the late 1950's. Since 1997, the index has risen about 83 percent.



Source: "Irrational Exuberance," 2nd Edition, 2006, by Robert J. Shiller

Bill Marsh/The New York Times

Procuring Tenders

- Find tenders in relevant areas
- Expensive, low quality services

GlobalTenders.com - A premium source of international business tenders, notices, national and international competitive bids, contracts, RFPs, RFQs.

GlobalTenders

Home / Categories / Search / Publish Tenders / Directory / Tenders

Website Marketing
Effective SEO
Improve Search Engine Ranking

SUSHI COUPONS*
in Oxford
*Example of an upcoming Deal

REGION-WISE TENDERS

- African Tenders
- American Tenders
- Asian Tenders
- Australian & Oceania
- Caribbean Tenders

Global Tenders - world's latest tender information

Global Tenders has one of the largest

All Tenders Live Tenders Closed Tenders

Within Search : --Select-- Search

No.	Tender ID & Tender Brief	State	Tender Value	Due Date
1	TR 2104776 (I) Improvements to Road from Tayakanahalli to Chowtayanahatti (KM 0.00 to 2.50) II) Improvements to Road from Kadekola to Nelabommanahalli (KM 0.00 to 2.20) in Kudligi Taluk. (Package No.: BLR-KDL-I)	KARNATAKA	1,21,00,000.00	15/11/2011
2	TR 2386052 Repair and Maintenance of Main Road from Badam Pahar Corner to Jhamadih CISF Colony Gate at CTPS.			
3	TR 2377019 Request for Qualification for Short-Listing of Concessionaire for Development and Operation of Basti-Mehndawal-Tamkuhi Road Section on PPP basis.			
4	TR 2391770 Improvement and Widening to Shetphal Medha Papans Road MDR-18 km 0/00 to 28/00 Tal. Madha.			
5	TR 2396806 Improvements to Sangola-Pandharpur-Shetphal-Kurduwadi-Paranda Rd.Km. 54/00 to 55/500 (Pro.BM,Carpet,Sealcoat)Tal.Mohol Dist.Solapur.			
6	TR 2396810 Improvements to Pandharpur-Shetphal-Kurduwadi Rd.Km.63/00 to 65/500(Pro.Carpet,sealcoat).			
7	TR 2396815 Improvements to Pandharpur Shetphal Kurduwadi Road Km.54/00 to 57/00 (Pro.Carpet,Sealcoat)Tal.Mohol Dist.Solapur.			
8	TR 2396817 Improvements to Pandharpur-Kurul Road Km.24/00 to 25/700			

International Tenders

Start today to get instant access to the latest [International Tenders](#).

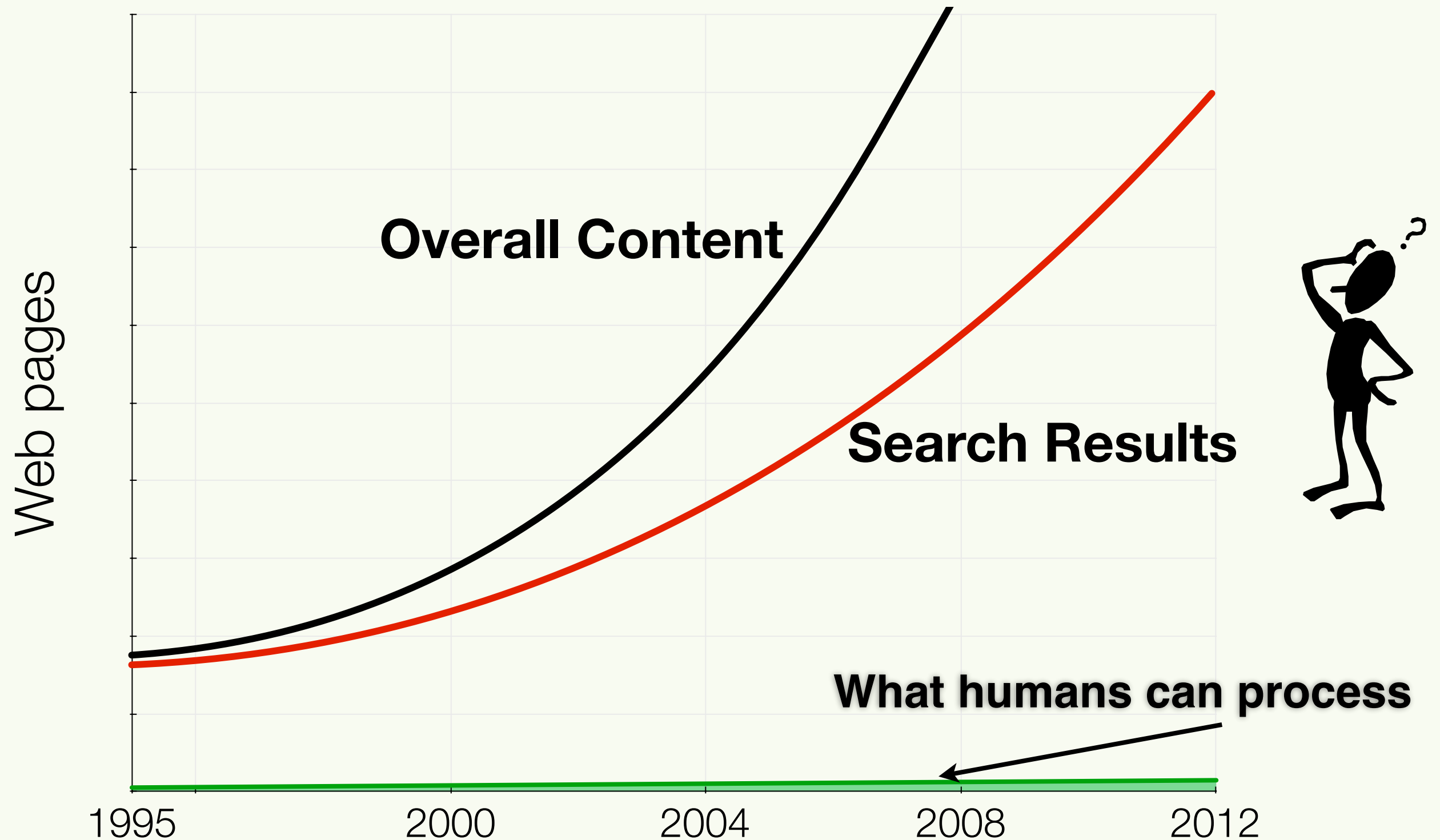
We deliver thousands of live Tender Notices every month. You'll receive regular Tenders from a wide range of companies and public sector organisations.

Start to enjoy this invaluable source of new sales opportunities today with our [4 weeks FREE Trial](#).

FREE Trial for 4 weeks
[click here](#)

CONTINUE

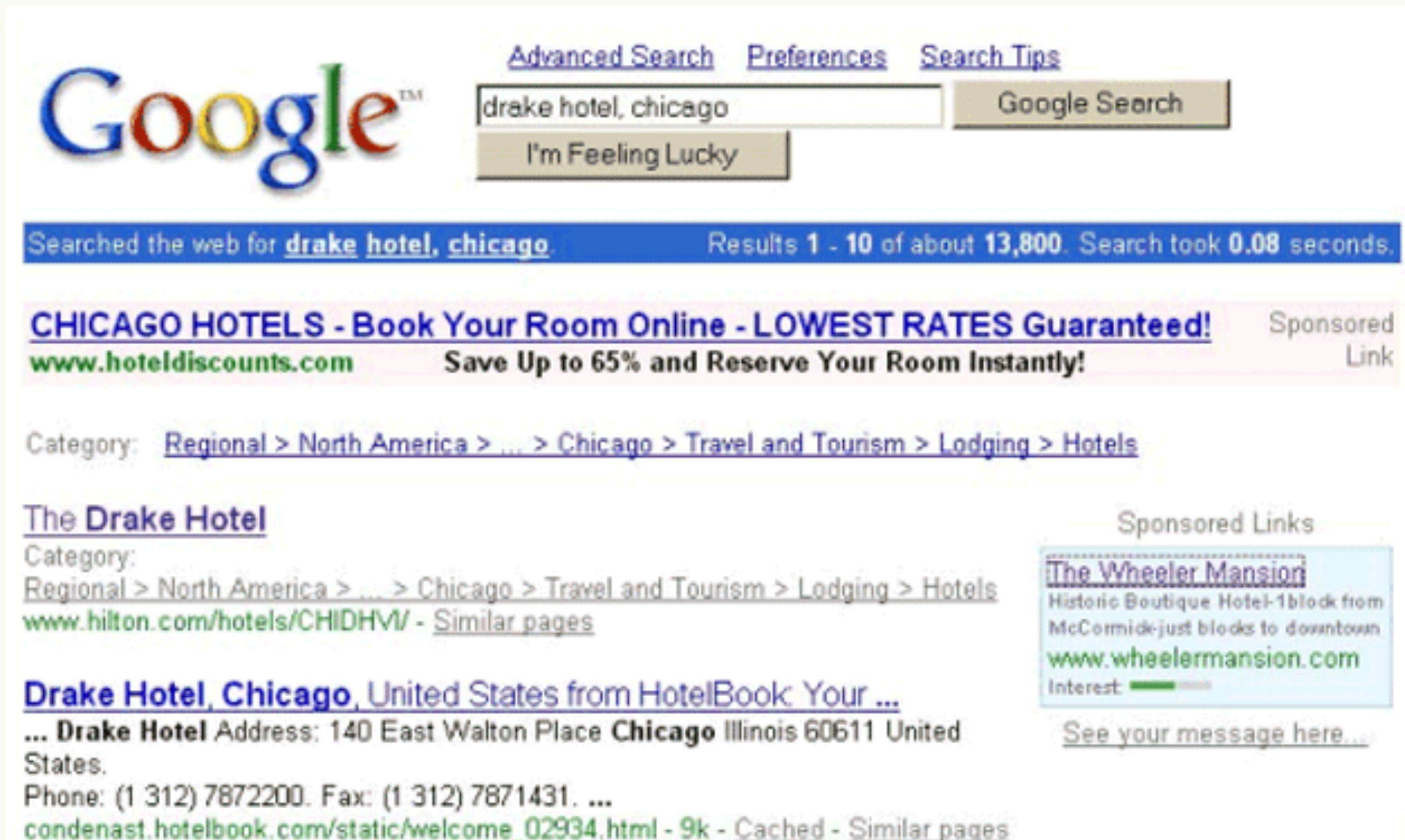
Search



Search

What's the biggest change to Google's results in the last decade?

Search



The image is a screenshot of a Google search interface from the year 2000. At the top, the Google logo is on the left, and navigation links for 'Advanced Search', 'Preferences', and 'Search Tips' are on the right. Below the logo is a search input field containing the text 'drake hotel, chicago'. To the right of the input field is a 'Google Search' button, and below it is an 'I'm Feeling Lucky' button. A blue banner below the search bar indicates the search was performed for 'drake hotel, chicago' and shows 'Results 1 - 10 of about 13,800. Search took 0.08 seconds.'

The first search result is a sponsored link for 'CHICAGO HOTELS - Book Your Room Online - LOWEST RATES Guaranteed!' from 'www.hoteldiscounts.com'. The text promises to 'Save Up to 65% and Reserve Your Room Instantly!'. To the right of this link is the text 'Sponsored Link'.

Below the sponsored link is a category breadcrumb: 'Category: [Regional](#) > [North America](#) > ... > [Chicago](#) > [Travel and Tourism](#) > [Lodging](#) > [Hotels](#)'.

The first organic search result is for 'The Drake Hotel'. It includes the same category breadcrumb and a link to 'www.hilton.com/hotels/CHIDHVI/' with a 'Similar pages' link.

The second organic search result is for 'Drake Hotel, Chicago, United States from HotelBook: Your ...'. It provides the address '140 East Walton Place Chicago Illinois 60611 United States', phone '(1 312) 7872200', fax '(1 312) 7871431', and a link to 'condenast.hotelbook.com/static/welcome_02934.html' with a '9k' file size, 'Cached' status, and 'Similar pages' link.

On the right side of the page, under the heading 'Sponsored Links', there is a box for 'The Wheeler Mansion'. It describes it as a 'Historic Boutique Hotel-1block from McCormick-just blocks to downtown' and provides the website 'www.wheelermansion.com'. Below this is an 'Interest' bar and a link to 'See your message here...'.

2000

Search

2019

Web Maps Images Videos News More Search tools

About 2,740,000 results (0.32 seconds)

The Drake Hotel
www.thedrakehotel.com/
 No other **Chicago** hotel has inspired more legendary acclaim than The Drake. Fine dining, Lake Michigan views, **Chicago's** Gold Coast neighborhood access.
 4.2 ★★★★★ 328 Google reviews · Write a review · £129▼

140 E Walton Pl, Chicago, IL 60611, United States
 +1 312-787-2200
 Accommodations · Palm Court · Packages

The Drake Hotel (Chicago, IL) - Hotel Reviews - TripAdvisor
www.tripadvisor.co.uk > ... > Illinois (IL) > Chicago > Chicago hotels
 ★★★★★ Rating: 4 - 1,391 reviews - Price range: ££
 The Drake Hotel, Chicago: See 1391 traveller reviews, 323 candid photos, and great deals for The Drake Hotel, ranked #50 of 166 hotels in Chicago and rated 4 ...

Drake Hotel (Chicago) - Wikipedia, the free encyclopedia
[en.wikipedia.org/wiki/Drake_Hotel_\(Chicago\)](http://en.wikipedia.org/wiki/Drake_Hotel_(Chicago))
 The Drake Hotel, 140 East Walton Place, **Chicago**, Illinois, is a luxury, full-service hotel, located downtown on the lake side of Michigan Avenue two blocks north ...

Magnificent Mile hotel Downtown Chicago - The Drake Hotel - Hilton
www3.hilton.com/en/hotels/illinois/the-drake-hotel.../index.html
 The Drake Hotel, a Magnificent Mile hotel in downtown **Chicago**, offers historic luxury **Chicago** hotel accommodations, first class service and diverse dining.

The Drake Hotel, Chicago, Illinois | Hotels.com
uk.hotels.com > United States Hotels > Illinois Hotels > Chicago Hotels
 Location. The Drake Hotel is located in downtown **Chicago**, adjacent to Michigan Avenue's renowned shops and restaurants. This property stands three blocks ...

The Drake Hotel in Chicago: Hotel Rates & Reviews on Orbitz
www.orbitz.com > Hotels > United States > Illinois > Chicago Hotels
 ★★★★★ Rating: 4.3 - 172 reviews
 The Drake Hotel, **Chicago**, Illinois: Find the best deals with user reviews, photos, and discount rates for The Drake Hotel at Orbitz. Get our lowest rates or cash ...

The Drake Hotel (Chicago, United States of America) | Expedia
www.expedia.co.uk > ... > Illinois > Chicago Hotels
 ★★★★★ Rating: 4.5 - 1,579 reviews
 The Drake Hotel - read reviews, look at the photos and get great deals. Book the The Drake Hotel with Expedia now and save!

Images

See photos See inside

The Drake Hotel
 £129 Book Directions Write a review

Address: 140 E Walton Pl, Chicago, IL 60611, United States
 Phone: +1 312-787-2200
 Hotel class: 5 stars

Reviews
 4.2 ★★★★★ 328 Google reviews

More reviews: priceline.com, insiderpages.com, blackbookmag.com, kudzu.com

People also search for

Four Seasons Hotel Chicago The Peninsula Chicago Omni Chicago Hotel Waldorf Astoria Chicago Hotel Chicago Marriott Downtown Magnificent Mile

Are you the business owner? Feedback

Location

Snippets

Knowledge graph

Reviews

CONCLUSION:

Search is becoming **vertical** to combat information **overload**.

What is data extraction?



Goal



~~Structured API
(XML/JSON)~~

~~Semantic API
(RDF)~~



.....>
template

HTML
interface

The image shows a real estate search interface with a purple background. At the top, there are tabs for 'Search', 'For sale', 'To rent', 'Research', 'Current values', and 'Sold prices'. Below these is a search bar containing the text 'e.g. Oxford or NW3'. There are several filter sections: 'Min price' and 'Max price' both set to 'No min' and 'No max' respectively; 'Type' set to 'Show all'; 'Bedrooms' set to 'No min'; 'Distance from location' set to 'This area only'; 'Added' set to 'Any time'; and 'Sort by' set to 'Highest price'. There is a 'Keywords' input field. Under the 'Include' section, there are four checkboxes: 'New homes', 'Retirement homes', 'Shared ownership homes', and 'Under offer or sold STC', all of which are checked. A red 'Search for sale' button is located at the bottom right. At the very bottom, there is a link that says 'Browse by area'.



HTML
interface

Search For sale To rent Research Current values Sold prices

e.g. Oxford or NW3

Min price No min Max price No max Type Show all Bedrooms No min Fewer options

Distance from location This area only

Added Any time Sort by Highest price

Keywords

Include

- ☒ New homes
- ☒ Retirement homes
- ☒ Shared ownership homes
- ☒ Under offer or sold STC

Search for sale

Browse by area

Form filling

WHITE WALLS people not just property

Home About us Sales Lettings/Management Free valuation EPC Careers Links Contact

Property in Oxford search results

« Go back to postcode area map 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 NEXT 1 -

White Road, East Hendred £1,499,000 New price

Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension. The five [...] [Read more](#)

Bedrooms: 15 | Bathrooms: 13 | Area: East Hendred

[View all 10 photos](#)

[View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

Bagley Wood Rd, Kennington OX1 £1,295,000

A substantial detached Edwardian family house with scope for updating and improvement, set within delightful gardens and wooded grounds extending to over five acres located on the edge of Bagley Wood. This fine [...] [Read more](#)

Bedrooms: 6 | Bathrooms: 3 | Area: Kennington

[View all 16 photos](#)

[View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)


[View all 16 photos](#)

Bagley Wood Rd, Kennington OX1

£1,295,000

A substantial detached Edwardian family house with scope for updating and improvement, set within delightful gardens and wooded grounds extending to over five acres located on the edge of Bagley Wood. This fine [...] [Read more](#)

Sold ✓

Bedrooms: 6 | Bathrooms: 3 | Area: Kennington

[View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

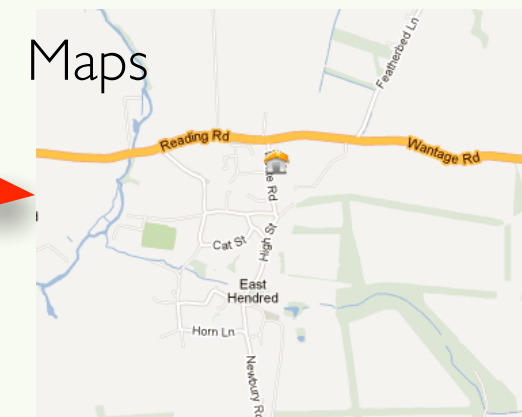
Object
identification

White Road, East Hendred

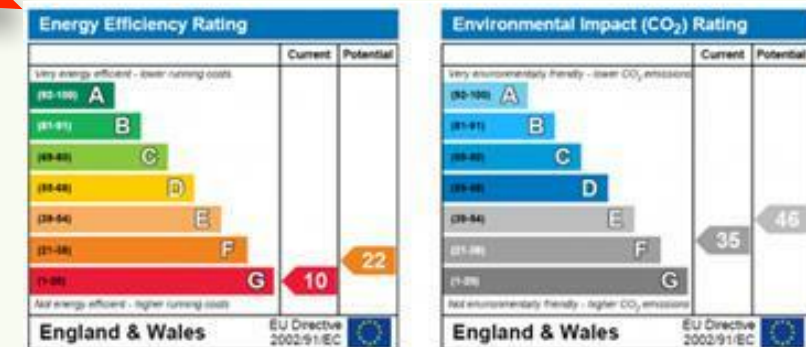
3
£1,499,950

[Arrange a viewing](#)
[Image gallery](#)
[360° virtual tour](#)
[3D Floorplan](#)
[Aerial view](#)
[Map view](#)
[Printable brochure](#)
[Email a friend](#)
[EPC](#)
[Free mortgage advice](#)


Maps



Energy Performance Chart



0% Coming soon to OX2

[Pre-register now](#)

Sales or Lettings: **Sales** Property Type: **House** Postcode Area: **OX12**
 Bedrooms: **15** Bathrooms: **13** Furnishing: **N/A**
 Area: **East Hendred** Tenure: **Freehold** Sq feet: **6648**

Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension.

The five bedroom family house benefits from generously proportioned living space with modern interiors, conservatory, dining room with feature bow window and stunning first floor views over substantial private gardens. This is further accompanied an attached recently purpose built AA 4* rated ten bedroom guest house/bed and breakfast wing.

In addition, subject to planning, further development potential is possible.

The property is situated within its own grounds of approximately 2 acres, Local transport links are available nearby, providing access to Oxford and Wantage. Road users have easy access to the A34 (M40) towards London.

Tables

Flat Text



Structured API
(XML/JSON)

Semantic API
(RDF)

HTML
interface

Form filling

WHITE WALLS
people not just property

Home About us Sales Lettings/Management Free valuation EPC Careers Links Contact us

Property in Oxford search results

Go back to postcode area map

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 NEXT 1 - 10 of properties

White Road, East Hendred £1,499,950

Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension. The five [...] [Read more](#)

Bedrooms: 15 | Bathrooms: 13 | Area: East Hendred

[View all 10 photos](#)

[View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

Bagley Wood Rd, Kennington OX1 £1,295,000

A substantial detached Edwardian family house with scope for updating and improvement, set within delightful gardens and wooded grounds extending to over five acres located on the edge of Bagley Wood. This fine [...] [Read more](#)

Bedrooms: 6 | Bathrooms: 3 | Area: Kennington

[View all 16 photos](#)

[View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

Domain database

Other providers



Cleaning &
integration

4

Other
Provider

Other
Provider

Other
Provider

Other
Provider

Object
identification

White Road, East Hendred 3 £1,499,950

[Arrange a viewing](#) [Image gallery](#) [360° virtual tour](#) [Floorplan](#) [Aerial view](#) [Map view](#) [Printable brochure](#) [Email a friend](#) [EPC](#) [Free mortgage advice](#)

Maps

Energy Performance Chart

Energy Efficiency Rating

Current	Potential
A	A
B	B
C	C
D	D
E	E
F	F
G	G

England & Wales EU Directive 2002/91/EC

Environmental Impact (CO₂) Rating

Current	Potential
A	A
B	B
C	C
D	D
E	E
F	F
G	G

England & Wales EU Directive 2002/91/EC

Tables

Sales or Lettings:	Property Type:	Postcode Area:
Sales	House	OX12
Bedrooms: 15	Bathrooms: 13	Furnishing: N/A
Area: East Hendred	Tenure: Freehold	Sq feet: 6648

Flat Text

Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension.

The five bedroom family house benefits from generously proportioned living space with modern interiors, conservatory, dining room with feature bow window and stunning first floor views over substantial private gardens. This is further accompanied an attached recently purpose built AA 4* rated ten bedroom guest house/bed and breakfast wing.

In addition, subject to planning, further development potential is possible.

The property is situated within its own grounds of approximately 2 acres, Local transport links are available nearby, providing access to Oxford and Wantage. Road users have easy access to the A34 (M40) towards London.

Coming soon to OX2

[Pre-register now](#)

Why is it hard?



Variety and volume

Thousands of sources; each different



Data quality

Different attributes, terminology, ...



Velocity: costly supervision

Difficult task; fast changing websites



Technically challenging

Ever changing target (HTML5, AJAX)

CONCLUSION:

- 1 Supervised techniques lack **scale**
- 2 Automatic techniques lack **accuracy**

Landscape

Manual

- Ad-hoc
- Scrapy & Pandas
- Very common

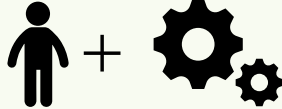
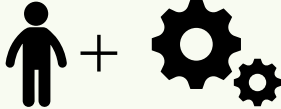
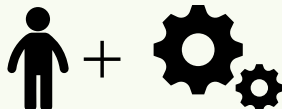
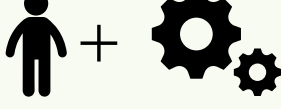
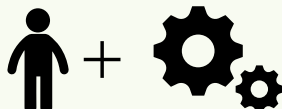
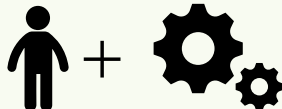
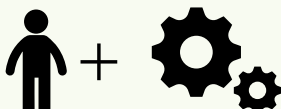
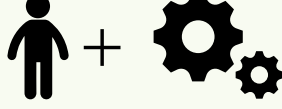
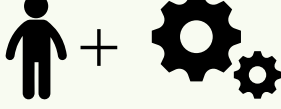
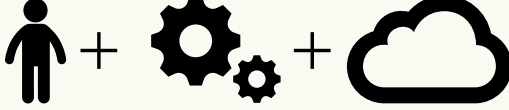
Automatic

- Fully algorithmic
- Active research
- First commercial products

Supervised

- Human + algorithm
- Most commercial products



	exploration	identification
scapinghub.com		
Mozenda		
Connotate		
promptcloud.com		
Lixto		
Blacklocus		
import.io		
diffbot.com		
DIADEM		



Need for Automatic Extraction

- Example: **Real Estate UK > 15000 sites**
 - many not covered by **aggregators**
 - list of all agencies easy to get (source discovery)
 - but: manual or **semi-automatic wrapping too expensive**
 - wrapper construction
 - testing
 - tracking changes
- **No existing tool or methodology** can do it fully automatically



Need for Automatic Extraction

- All **search engine providers** need it! Many work on it.
 - vertical search
 - object search
 - semantic search

no one really has done this successfully at scale yet

*current technologies are **not good enough yet** to provide what search engines really need. [...] any successful approach would probably need a **combination of knowledge and learning***

Alon Halevy, Google, Feb. 2009



Need for Automatic Extraction

An analysis of structured data on the web, Dalvi et al. (Yahoo) VLDB 2012

- This study shows: significant long-tail effect for many attributes
 - >1000 sites to get above 80% coverage required
- Examples of these attributes:
 - phone numbers and home pages of companies
 - restaurants, car sellers, hotels, banks, ...
 - ISBN of books
 - reviews of hotels and restaurants

*for many kinds of information one may have to extract
from thousands of sites in order to build a
 comprehensive database, even when we restrict to a
 given domain with known popular top sites*



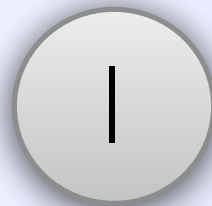
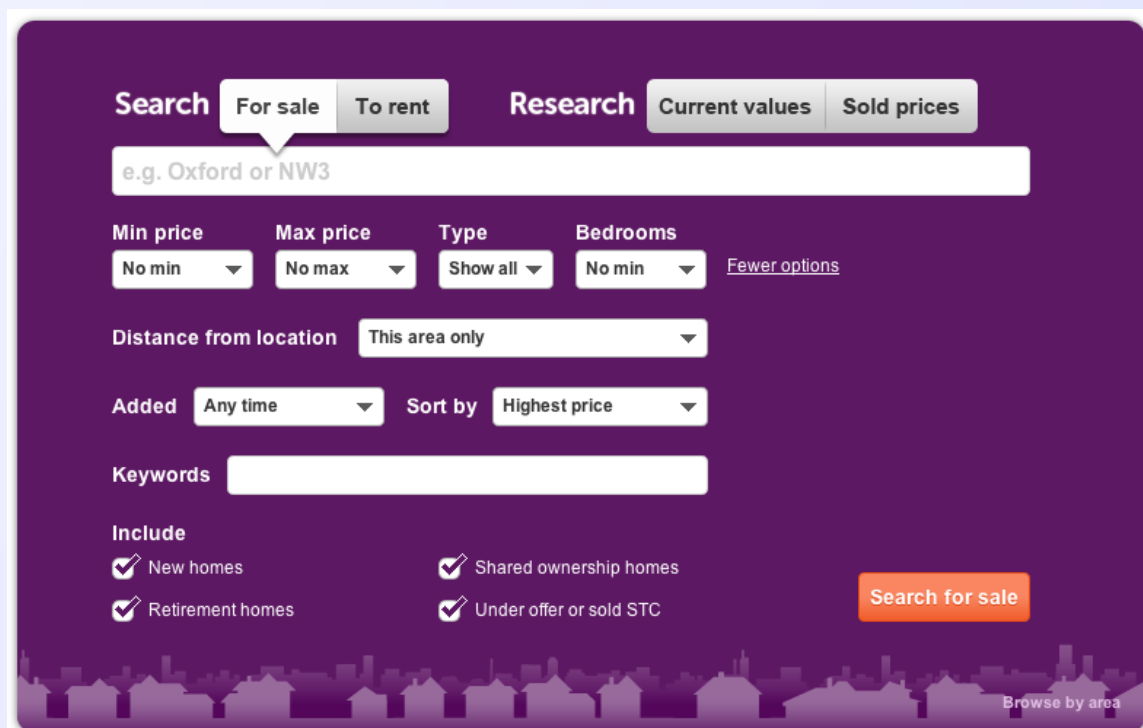
DIADEM | domain-centric intelligent automated
data extraction methodology





Form

Understanding & Filling



Search **For sale** **To rent** **Research** **Current values** **Sold prices**

e.g. Oxford or NW3

Min price: No min Max price: No max Type: Show all Bedrooms: No min [Fewer options](#)

Distance from location: This area only

Added: Any time Sort by: Highest price

Keywords:

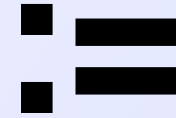
Include

☒ New homes ☒ Shared ownership homes

☒ Retirement homes ☒ Under offer or sold STC

Search for sale

Browse by area

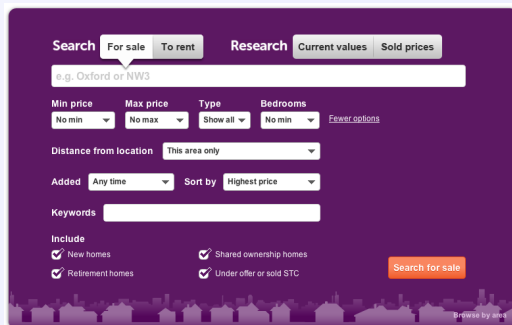


Form
Understanding & Filling

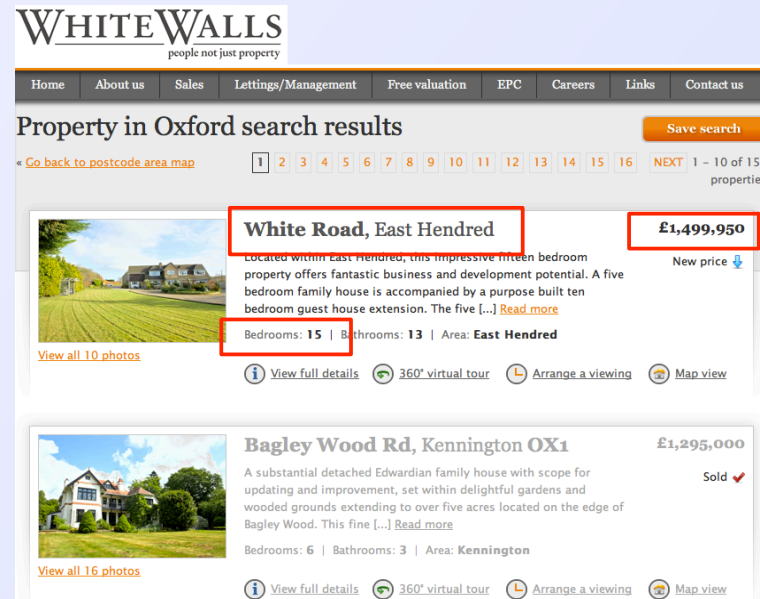
Object
identification & alignment

1

2



Result pages



White Road, East Hendred £1,499,950

Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension. The five [...] [Read more](#)

Bedrooms: 15 | Bathrooms: 13 | Area: East Hendred

[View all 10 photos](#) [View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

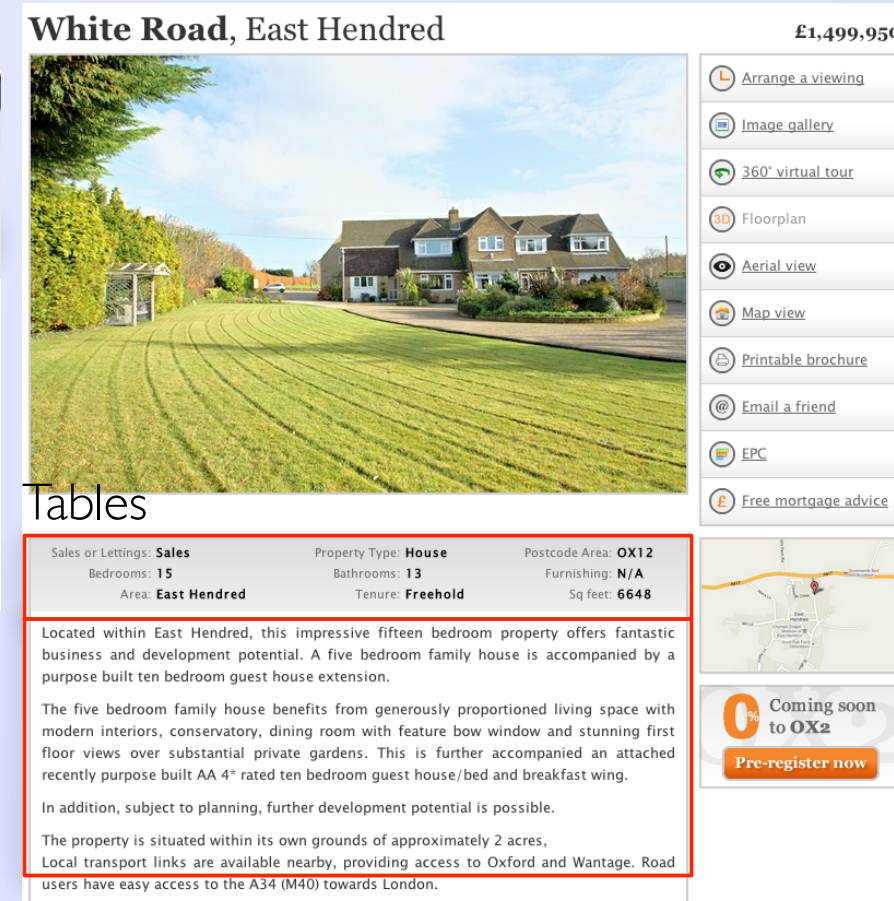
Bagley Wood Rd, Kennington OX1 £1,295,000

A substantial detached Edwardian family house with scope for updating and improvement, set within delightful gardens and wooded grounds extending to over five acres located on the edge of Bagley Wood. This fine [...] [Read more](#)

Bedrooms: 6 | Bathrooms: 3 | Area: Kennington

[View all 16 photos](#) [View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

Single entity (details) pages



White Road, East Hendred £1,499,950

[Arrange a viewing](#) [Image gallery](#) [360° virtual tour](#) [3D Floorplan](#) [Aerial view](#) [Map view](#) [Printable brochure](#) [Email a friend](#) [EPC](#) [Free mortgage advice](#)

Tables

Sales or Lettings: Sales	Property Type: House	Postcode Area: OX12
Bedrooms: 15	Bathrooms: 13	Furnishing: N/A
Area: East Hendred	Tenure: Freehold	Sq feet: 6648

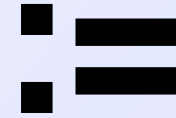
Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension.

The five bedroom family house benefits from generously proportioned living space with modern interiors, conservatory, dining room with feature bow window and stunning first floor views over substantial private gardens. This is further accompanied an attached recently purpose built AA 4* rated ten bedroom guest house/bed and breakfast wing.

In addition, subject to planning, further development potential is possible.

The property is situated within its own grounds of approximately 2 acres, Local transport links are available nearby, providing access to Oxford and Wantage. Road users have easy access to the A34 (M40) towards London.

Coming soon to OX2
[Pre-register now](#)



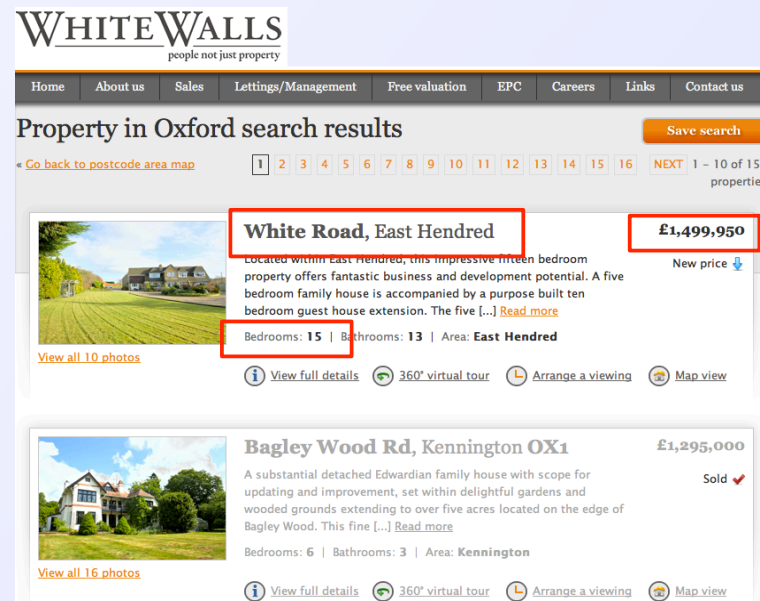
Object
identification & alignment

Context-driven
block analysis

2

3

Result pages



White Walls
people not just property

Property in Oxford search results

Go back to postcode area map

White Road, East Hendred £1,499,950

Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension. The five [...] [Read more](#)

Bedrooms: 15 | Bathrooms: 13 | Area: East Hendred

[View all 10 photos](#) [View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

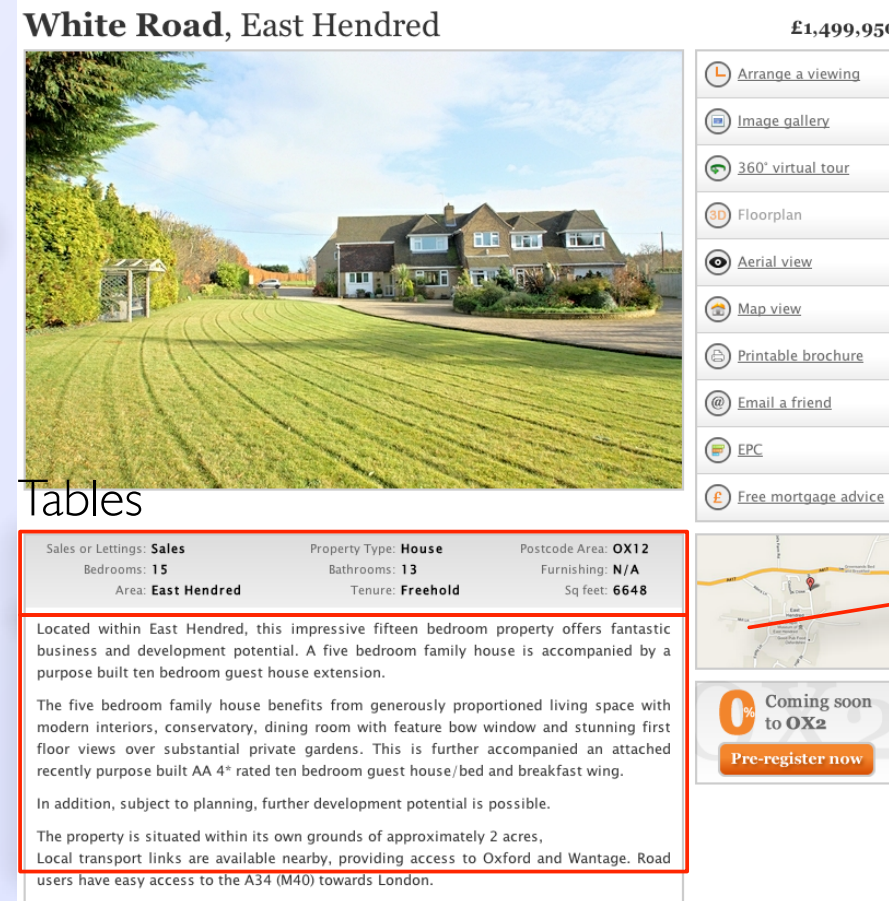
Bagley Wood Rd, Kennington OX1 £1,295,000

A substantial detached Edwardian family house with scope for updating and improvement, set within delightful gardens and wooded grounds extending to over five acres located on the edge of Bagley Wood. This fine [...] [Read more](#)

Bedrooms: 6 | Bathrooms: 3 | Area: Kennington

[View all 16 photos](#) [View full details](#) [360° virtual tour](#) [Arrange a viewing](#) [Map view](#)

Single entity (details) pages



White Road, East Hendred £1,499,950

[Arrange a viewing](#) [Image gallery](#) [360° virtual tour](#) [Floorplan](#) [Aerial view](#) [Map view](#) [Printable brochure](#) [Email a friend](#) [EPC](#) [Free mortgage advice](#)

Tables

Sales or Lettings: Sales	Property Type: House	Postcode Area: OX12
Bedrooms: 15	Bathrooms: 13	Furnishing: N/A
Area: East Hendred	Tenure: Freehold	Sq feet: 6648

Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension.

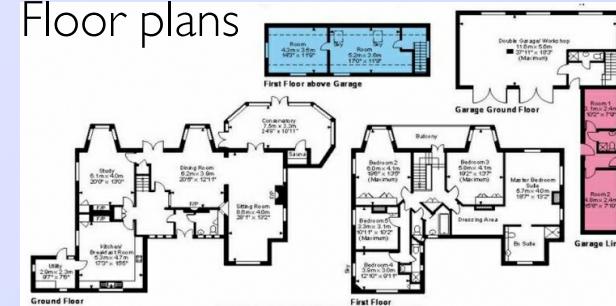
The five bedroom family house benefits from generously proportioned living space with modern interiors, conservatory, dining room with feature bow window and stunning first floor views over substantial private gardens. This is further accompanied an attached recently purpose built AA 4* rated ten bedroom guest house/bed and breakfast wing.

In addition, subject to planning, further development potential is possible.

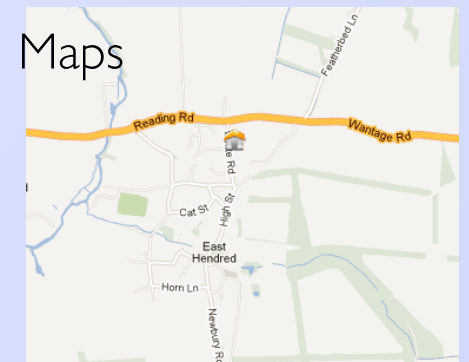
The property is situated within its own grounds of approximately 2 acres, Local transport links are available nearby, providing access to Oxford and Wantage. Road users have easy access to the A34 (M40) towards London.

Coming soon to OX2
[Pre-register now](#)

Floor plans

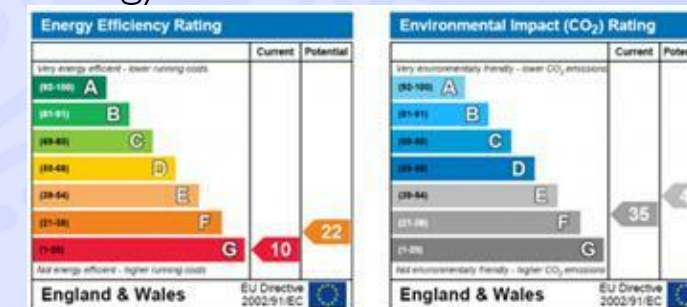


Maps



Flat Text

Energy Performance Chart





Identification & alignment

Context-driven **block** analysis

Context-driven **block** analysis

XPath Wrapper

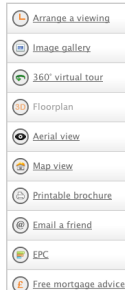
XPath Wrapper

Data integration

es

White Road, East Hendred

£1,499,950



Tables

Sales or Lettings: Sales	Property Type: House	Postcode Area: OX12
Bedrooms: 15	Bathrooms: 10	Furnishing: N/A
Area: East Hendred	Tenure: Freehold	Sq feet: 6648

Located within East Hendred, this impressive five bedroom family house property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension.

The five bedroom family house benefits from generously proportioned living space with modern interiors, conservatory, dining room with feature bow window and stunning first floor views over substantial private gardens. This is further accompanied an attached recently purpose built AA 4* rated ten bedroom guest house/bed and breakfast wing.

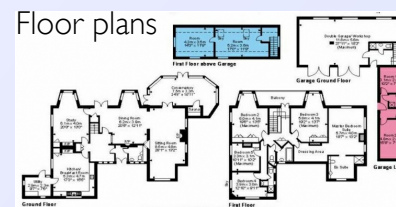
In addition, subject to planning, further development potential is possible.

The property is situated within its own grounds of approximately 2 acres.

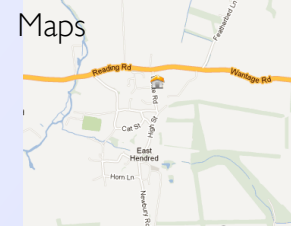
Local transport links are available nearby, providing access to Oxford and Watlington. Road users have easy access to the A34 (M40) towards London.

Flat Text

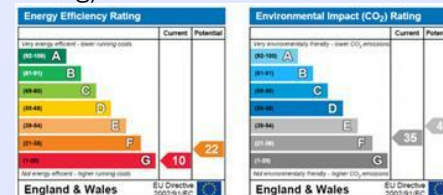
Floor plans



Maps



Energy Performance Chart





Form
Understanding & Filling

1

Object
identification & alignment

2

Context-driven
block analysis

3

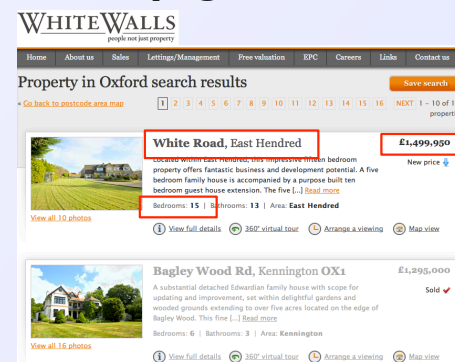
OXPath
Wrapper

4

Cloud extraction

Data integration

Result pages



Property in Oxford search results

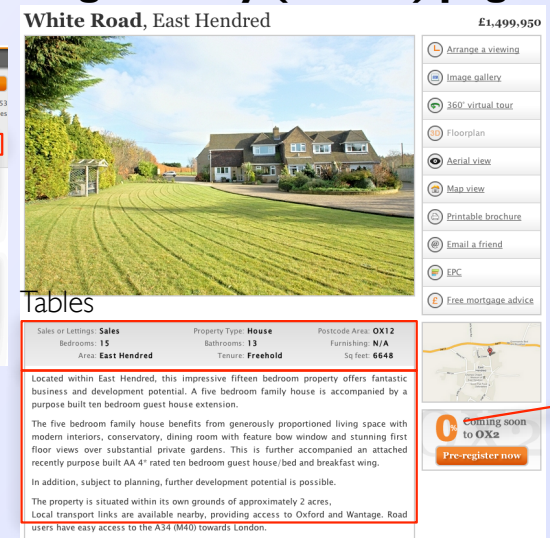
White Road, East Hendred £1,499,950

Bedrooms: 15 | Bathrooms: 13 | Area: East Hendred

Bagley Wood Rd, Kennington OX1 £1,095,000

Bedrooms: 6 | Bathrooms: 3 | Area: Kennington

Single entity (details) pages



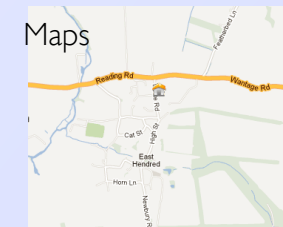
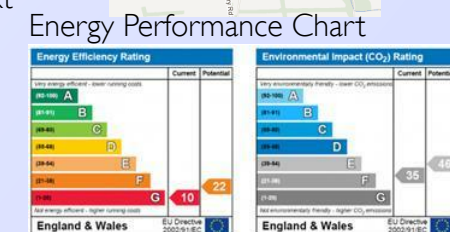
White Road, East Hendred £1,499,950

Arrange a viewing
Image gallery
360° virtual tour
Floorplan
Aerial view
Map view
Printable brochure
Email a friend
EPC
Free mortgage advice

Tables

Sales or Lettings:	Sales	Property Type:	House	Postcode Area:	OX12
Bedrooms:	15	Bathrooms:	13	Furnishing:	N/A
Area:	East Hendred	Tenure:	Freehold	Sq feet:	6648

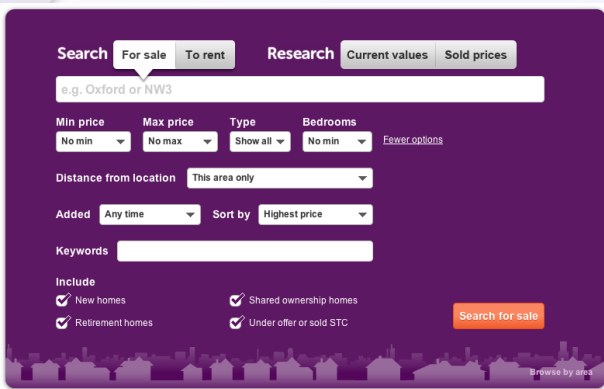
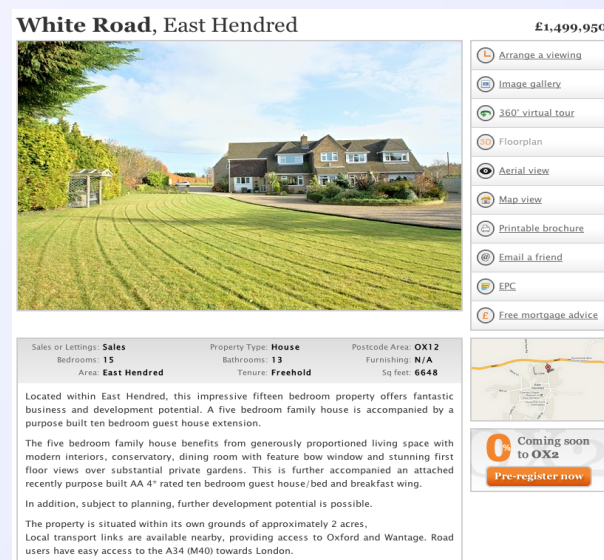
Located within East Hendred, this impressive fifteen bedroom property offers fantastic business and development potential. A five bedroom family house is accompanied by a purpose built ten bedroom guest house extension. The five bedroom family house benefits from generously proportioned living space with modern interiors, conservatory, dining room with feature bow window and stunning first floor views over substantial private gardens. This is further accompanied an attached recently purpose built AA 4* rated ten bedroom guest house/bed and breakfast wing. In addition, subject to planning, further development potential is possible. The property is situated within its own grounds of approximately 2 acres. Local transport links are available nearby, providing access to Oxford and Wantage. Road users have easy access to the A34 (M40) towards London.



5

Exploration Control
and Integration Language

HTML page



Detectors

Text
annotations

HTML
rendering

Document
structure

Flat text
dependencies

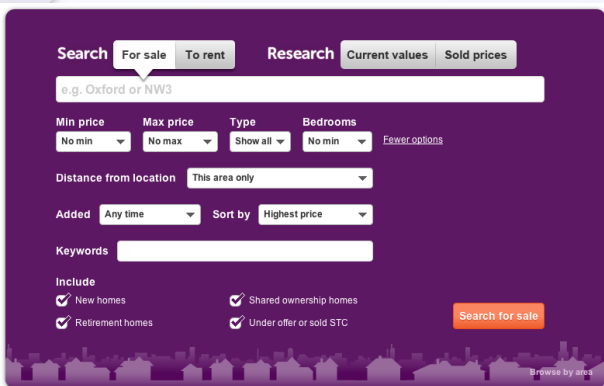
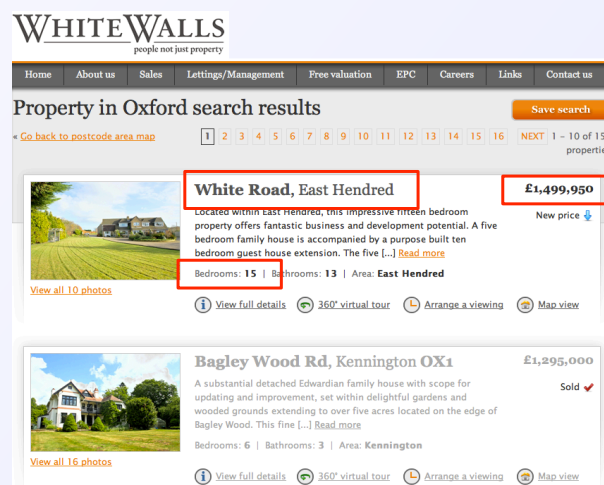
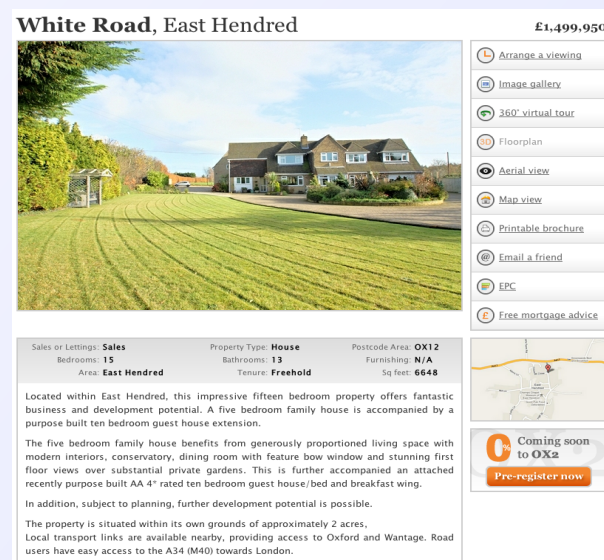
Image
features

URL
analysis

...

Observed phenomena

HTML page



Detectors

Text
annotations

HTML
rendering

Document
structure

Flat text
dependencies

Image
features

URL
analysis

...

Block
classification

Table
understanding

EPC chart
understanding

Floor plan
understanding

Observed phenomena

Inferred phenomena

Reasoners

Detectors

Text

annotations

HTML

rendering

Document

structure

Flat text

dependencies

Image

features

URL

analysis

...

Block

classification

Table

understanding

EPC chart

understanding

Floor plan

understanding

Form

Understanding

Object

identification & alignment

Context-driven
block analysis

Domain
knowledge



Observed phenomena

Inferred phenomena

Phenomenology

Concepts



Knowledge Graph Management

Location

Snippets

Web Maps Images Videos News More Search tools

About 2,740,000 results (0.32 seconds)

The Drake Hotel

www.thedrakehotel.com/

No other **Chicago** hotel has inspired more legendary acclaim than The Drake. Fine dining, Lake Michigan views, **Chicago's** Gold Coast neighborhood access.

4.2 ★★★★★ 328 Google reviews · Write a review · £129▼

140 E Walton Pl, Chicago, IL 60611, United States
+1 312-787-2200

Accommodations · Palm Court · Packages

The Drake Hotel (Chicago, IL) - Hotel Reviews - TripAdvisor

www.tripadvisor.co.uk > ... > Illinois (IL) > Chicago > Chicago hotels

★★★★★ Rating: 4 - 1,391 reviews - Price range: ££

The Drake Hotel, **Chicago**: See 1391 traveller reviews, 323 candid photos, and great deals for The Drake Hotel, ranked #50 of 166 hotels in Chicago and rated 4 ...

Drake Hotel (Chicago) - Wikipedia, the free encyclopedia

[en.wikipedia.org/wiki/Drake_Hotel_\(Chicago\)](http://en.wikipedia.org/wiki/Drake_Hotel_(Chicago))

The Drake Hotel, 140 East Walton Place, **Chicago**, Illinois, is a luxury, full-service hotel, located downtown on the lake side of Michigan Avenue two blocks north ...

Magnificent Mile hotel Downtown Chicago - The Drake Hotel - Hilton

www3.hilton.com/en/hotels/illinois/the-drake-hotel.../index.html

The Drake Hotel, a Magnificent Mile hotel in downtown **Chicago**, offers historic luxury **Chicago** hotel accommodations, first class service and diverse dining.

The Drake Hotel, Chicago, Illinois | Hotels.com

uk.hotels.com > United States Hotels > Illinois Hotels > Chicago Hotels

Location. The Drake Hotel is located in downtown **Chicago**, adjacent to Michigan Avenue's renowned shops and restaurants. This property stands three blocks ...

The Drake Hotel in Chicago: Hotel Rates & Reviews on Orbitz

www.orbitz.com > Hotels > United States > Illinois > Chicago Hotels

★★★★★ Rating: 4.3 - 172 reviews

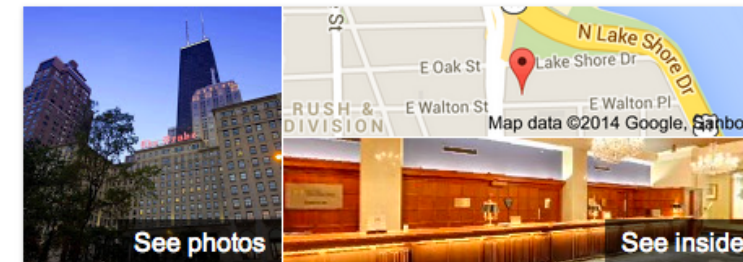
The Drake Hotel, **Chicago**, Illinois: Find the best deals with user reviews, photos, and discount rates for The Drake Hotel at Orbitz. Get our lowest rates or cash ...

The Drake Hotel (Chicago, United States of America) | Expedia

www.expedia.co.uk > ... > Illinois > Chicago Hotels

★★★★★ Rating: 4.5 - 1,579 reviews

The Drake Hotel - read reviews, look at the photos and get great deals. Book the The Drake Hotel with Expedia now and save!



The Drake Hotel

£129 Book ▼

Directions

Write a review

Address: 140 E Walton Pl, Chicago, IL 60611, United States

Phone: +1 312-787-2200

Hotel class: 5 stars

Reviews

4.2 ★★★★★ 328 Google reviews

More reviews: priceline.com, insiderpages.com, blackbookmag.com, kudzu.com

People also search for



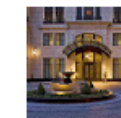
Four Seasons Hotel Chicago



The Peninsula Chicago



Omni Chicago Hotel



Waldorf Astoria Chicago Hotel



Chicago Marriott Downtown Magnificent Mile

Are you the business owner?

Feedback

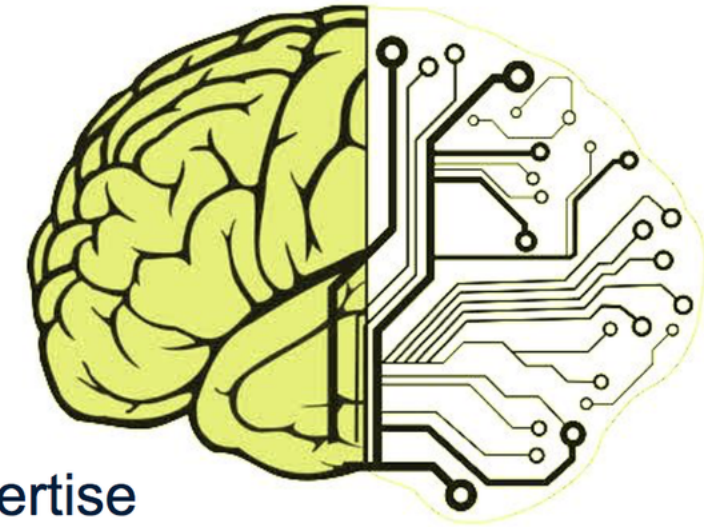
Knowledge graph

Reviews

The Future of Data Storage and Access: Knowledge Graphs

Next AI frontier: automating knowledge work

- 230 million knowledge workers worldwide
- Up to \$6 trillion in efficiency gains to be unlocked from automating knowledge-heavy business processes and subject-matter expertise
- Enterprises need a **knowledge graph platform** to encode human expertise and automate complex decision-making



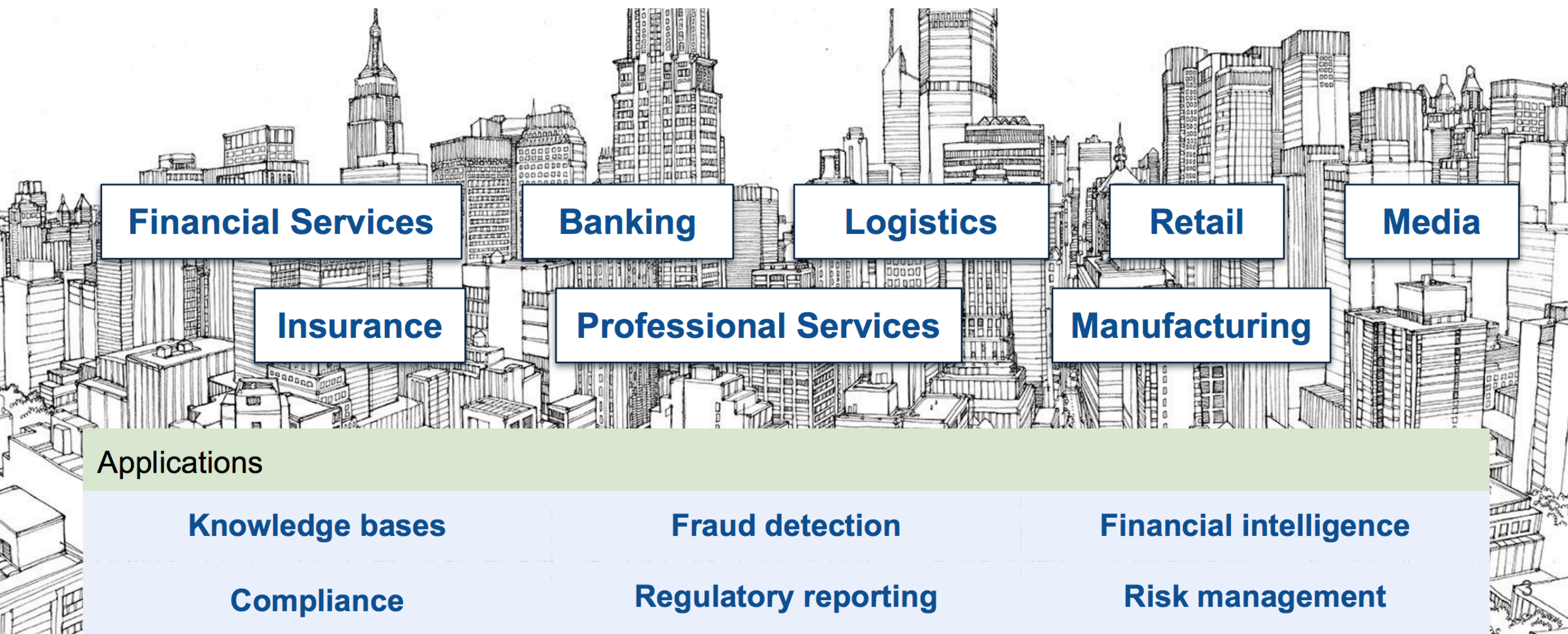
Operational Risk & Explainable AI

- Complex, **rapidly increasing obligations in compliance** and regulatory reporting: > 50,000 yearly updates in worldwide financial regulation
- Enterprises need to provide **audit trail** on how automated decision-making is made (e.g. **GDPR** gives EU consumers the “right to an explanation”)

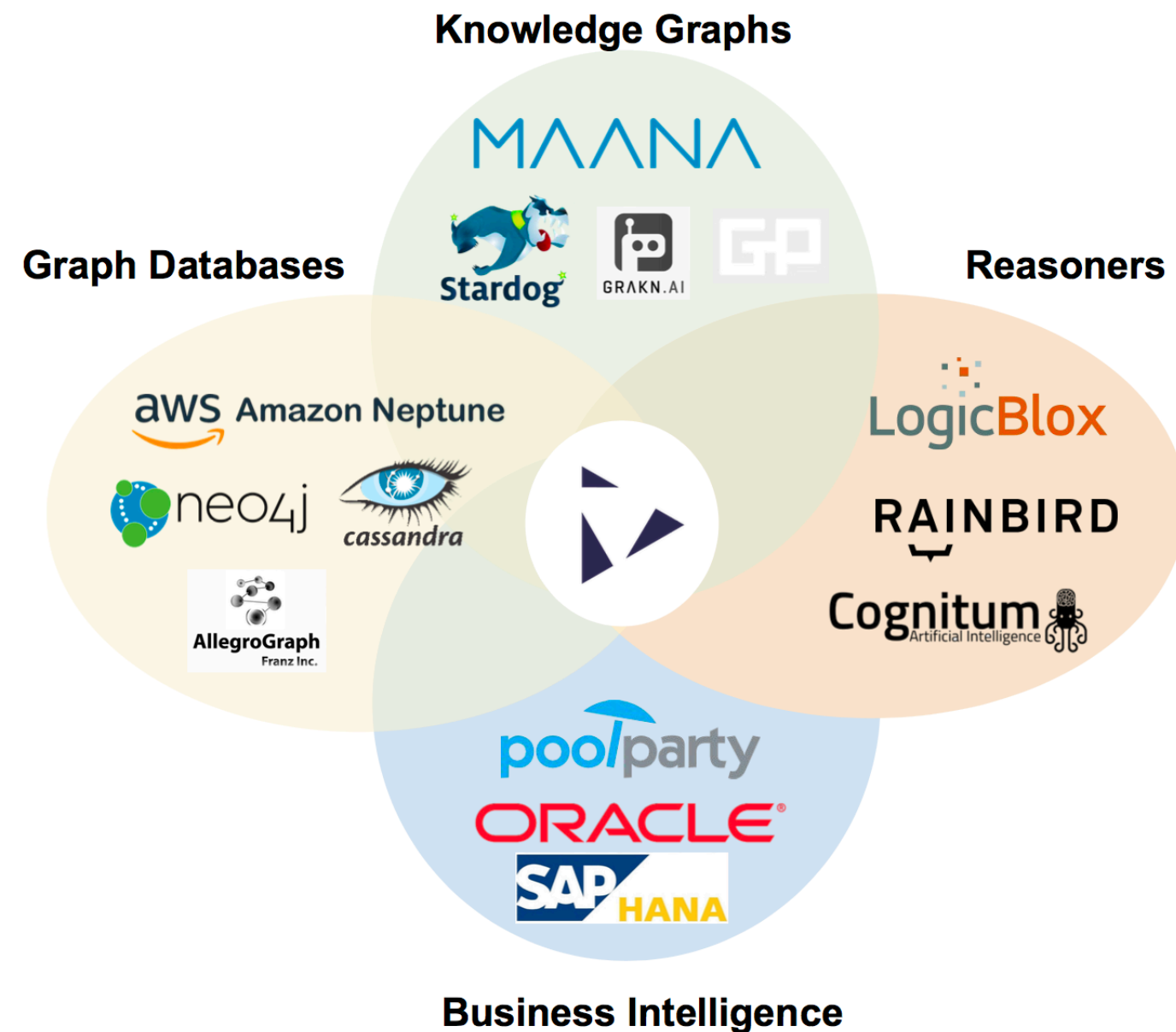
Urgent need for Knowledge Graph Management Systems (KGMS)

\$150 billion worldwide market for knowledge management, predictive analytics, regulatory and compliance software, with annual growth rate of 17.8%

Source: McKinsey



Urgent need for Knowledge Graph Management Systems (KGMS)

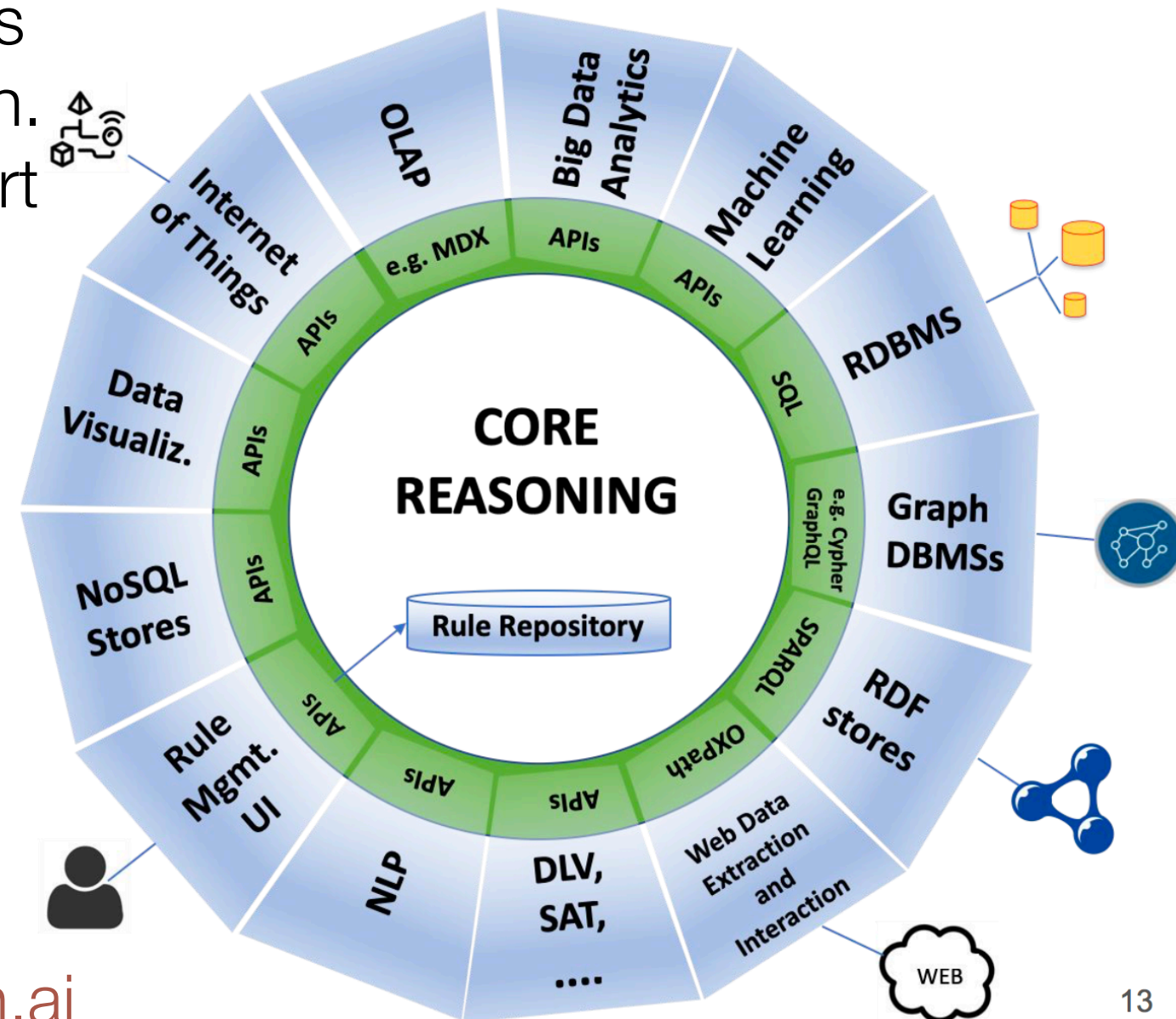


- Graph databases only provide **limited reasoning** over domain knowledge, and don't allow to encode & reuse expertise across domains
- Knowledge graph vendors **lock in into proprietary data formats**, with costly data migration
- Reasoners do not scale on large network data
- Machine Learning-based approaches are black-box, **non-compliant** due to lack of audit trail, and introduce decision biases

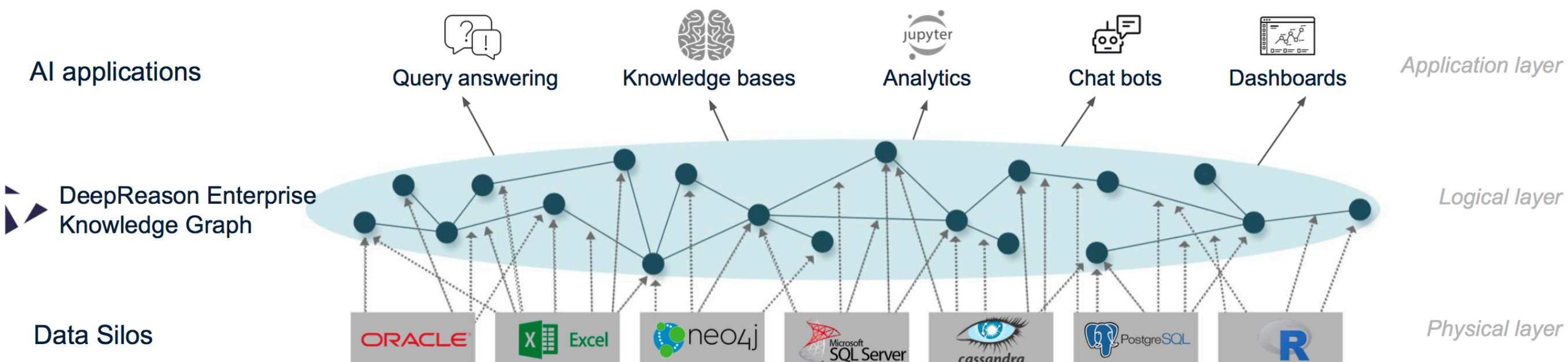
VADA project



- VADA project is a scientific collaboration between the Universities of Oxford, Manchester and Edinburgh.
- VADA aims to create a state of the art system for data wrangling and KG creation and management.
- A Reasoning Engine
- Database connectors
- APIs to Machine Learning, NLP, Constraint Solvers, Web Data Extraction
- A RESTful API
- A User Interface
- Commercial spinout — DeepReason.ai



DeepReason.ai's **Knowledge Graph** lets enterprises automate knowledge workflows and build next-generation AI applications



It unifies knowledge across siloed enterprise data, **without data migration**

It provides a rich **AI language to reason over domain knowledge & business rules**

It provides an **audit trail** for decisions: contributing factors, data source, and certainty

It has an **open architecture**: integrates with Machine Learning and connects to ontologies & data workflows

Data Management in DeepReason.ai

- DeepReason.ai can reason on **any data schema (relational, document and graph databases)** in contrast to most competitors which impose data migration (into RDF or property graphs).
- DeepReason.ai works as **in-memory virtual knowledge graph**, on a streaming basis with distributed caching.
- If required, it can use a **privileged materialised graph** – such as JanusGraph, Cassandra, DGraph – to store Knowledge Graph facts.

ORACLE®



Pure ML can often fail: creditworthiness example



up to £10,000



£8,500



£12,000



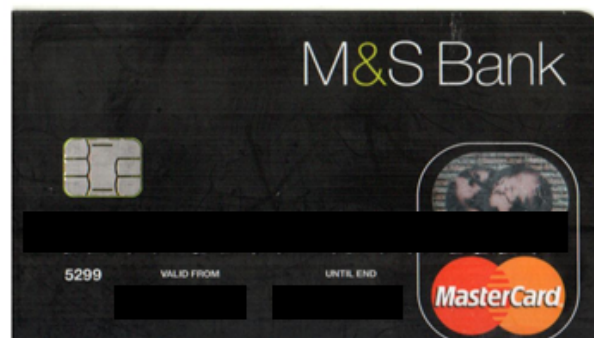
up to EUR 10,000



up to EUR 20,000



£500



£ 8,000



£ 12,500



EUR 14,000

Explanation

A machine-learning program has “reasonably” learned:

People who live in a joint household with someone who does not pay their bills are likely to fail repaying their own debts.

This ethically questionable rule was applied to wrong data.

Explanation

A machine-learning program has “reasonably” learned:

People who live in a joint household with someone who does not pay their bills are likely to fail repaying their own debts.

This ethically questionable rule was applied to wrong data.

A human credit rating expert would instead use of the rule:

If property owners move into their recently bought one-family property, then the previous occupiers have most likely moved out.

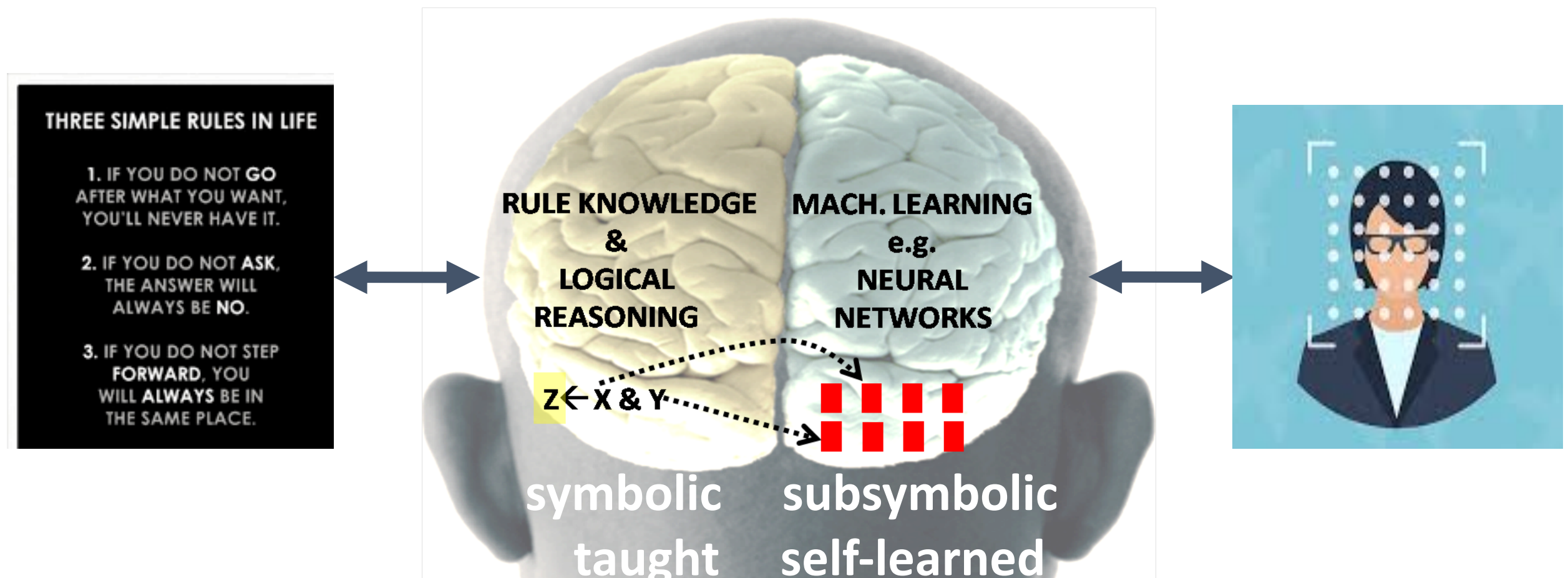
This rule can be used to update the database before applying machine learning.

Knowledge Graph Management Systems (KGMS)

KGMS combine the power of rule-based reasoning with machine learning over Big Data:

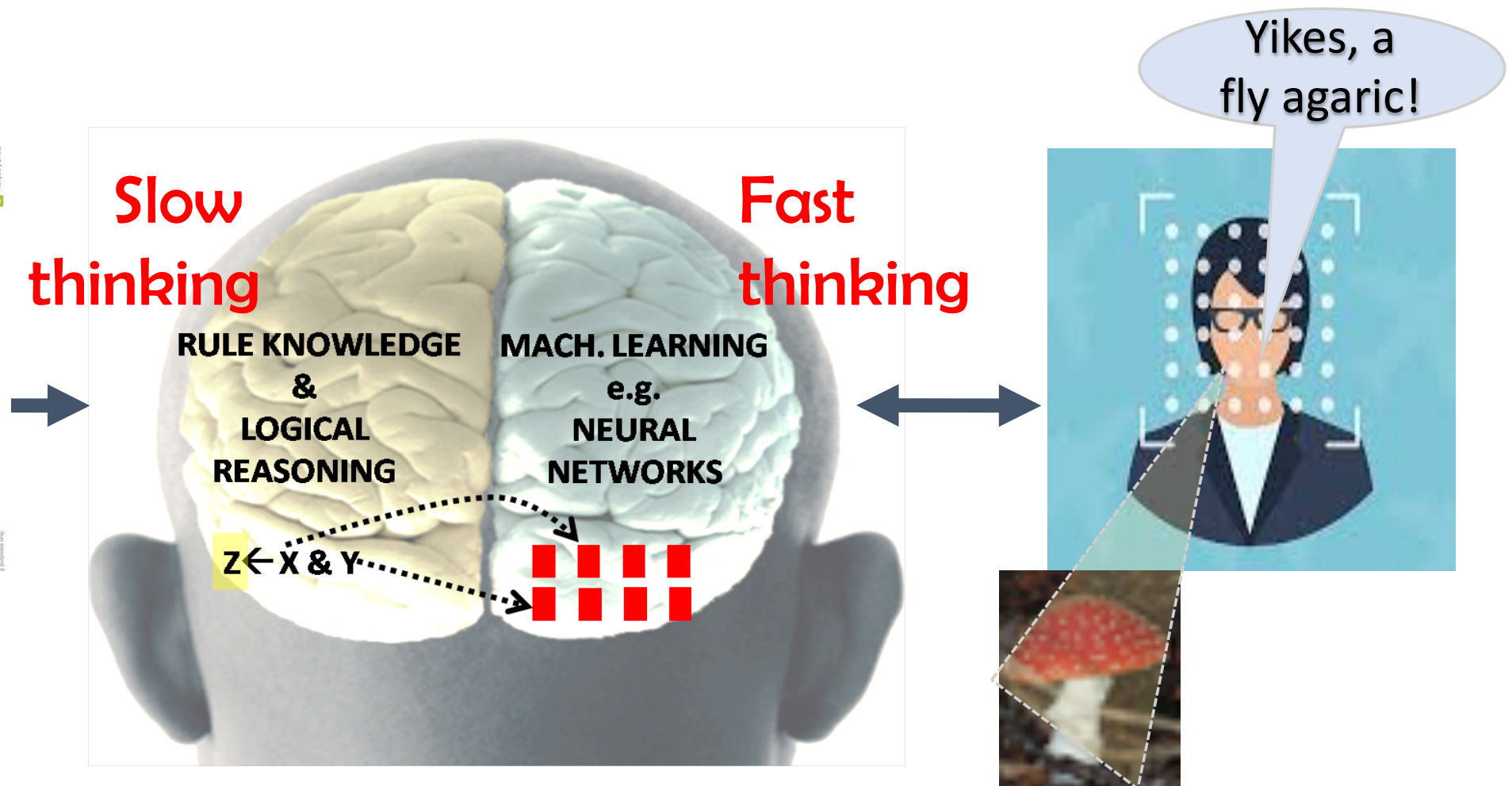
$$\text{KGMS} = \text{KBMS} + \text{Big Data} + \text{Analytics}$$

Misusing the lateralization thesis for illustration





Grandma: "Fly agarics are poisonous mushrooms. If you eat a poisonous mushroom, you may die".



DeepReason.ai Integration with ML

DeepReason.ai provides a **tight integration between Logical Reasoning and Machine Learning**. It supports the most popular ML frameworks and offers **3 modes of interaction**:



ML for Feature Engineering

DeepReason prepares a relation as ML input.
ML classifies facts and sends them into the core reasoning system.

On-demand classification

ML package acts as a special predicate.
Called by DeepReason's core reasoning system.

Rule Mining

Rule mining: ML infers rules from data.
Rules are translated into probabilistic rules.

DeepReason.ai Integration with Web Data Extraction

DeepReason.ai provides **native integration with automated web data extraction and IoT**, using the rich querying language XPath accessible directly from its reasoning engine.

Example VADALOG binding, extracting unstructured company registry data from the open web

```
@qbind("Own", "xpath",  
      "doc('http://company_register.com/ownerships')  
      /descendant::field()[1]/{$1}  
      /following::a[.##='Search']/ {click/  
      /(//a[.##='Next']/ {click/})  
      //div[@class='c']: [./span[1]:][./span[3]:]")
```

The logo for XPath, featuring the word "xpath" in a lowercase, sans-serif font. The "x" is stylized with a small spider-like figure integrated into its structure.

Furche, T., Gottlob, G., Grasso, G., Schallhart, C., & Sellers, A.. VLDB Journal, 2013

DeepReason.ai IDE

DeepReason.ai offers a **complete IDE** to create, manage, evaluate, analyse and debug Vadalog rules. It is based on the popular Jupyter environment.

UI Features:

- Rule authoring, execution
- Interaction with Python and R
- Code analysis
- Explanations (Proof Trees and Audit Trails)
- Visualisation

The screenshot displays the DeepReason.ai IDE interface, which is based on the Jupyter environment. The interface includes a sidebar with a file explorer showing a project named 'DeepReason_Shared' with files like 'Gremlin', 'imgs', and '1. KG Enrichment.ipynb'. The main area shows a Jupyter notebook titled 'Company Control' with the following Vadalog rules:

```
Company Company1 controls company Company2 if Company1 owns directly or indirectly  
Company Company1 controls company Company2 if Company1 owns directly or indirectly  
In [19]: %%explain("controls("A","C").").  
controls(Company1, Company2) :- ownsIndirectly(Company1, Company2, Share), Share > 0.5.  
ownsIndirectly(Company1, Company2, Share12) :- ownsDirectly(Company1, Company2, Share),  
Share12 = msum(Share, <Company2>).  
ownsIndirectly(Company1, Company2, Share12) :- ownsIndirectly(Company1, Company3, Share13),  
ownsDirectly(Company3, Company2, Share32),  
Share12 = msum(Share13 * Share32, <Company3>).  
ownsDirectly("A", "B", 0.5).  
ownsDirectly("A", "C", 0.4).  
ownsDirectly("B", "C", 0.4).  
ownsDirectly("C", "D", 0.5).  
ownsDirectly("B", "D", 0.4).  
@output("controls").  
%%explain("controls("A","C").").
```

Below the code, a proof tree is visualized, showing the logical derivation of the goal `controls(A, C)`. The tree starts with the goal node `controls(A, C)`, which is linked to the rule `controls(A, C) :- ownsIndirectly(A, C, 0.6), 0.6 > 0.5`. This leads to the goal `ownsIndirectly(A, C, 0.6)`, which is further decomposed into `ownsIndirectly(A, C, 0.6) :- ownsIndirectly(A, B, 0.5), ownsDirectly(B, C, 0.4), 0.6 = msum(0.5 * 0.4, <B>)` and `ownsIndirectly(A, C, 0.6) :- ownsDirectly(A, C, 0.6)`. The first branch further decomposes into `ownsIndirectly(A, B, 0.5)` and `ownsDirectly(B, C, 0.4)`, which are then linked to their respective base rules.



DeepReason.ai Implementation

Knowledge engineering

- Domain knowledge and subject-matter expertise is translated into rules-based model by data scientists.
- DeepReason.ai provides this in the initial pilot work, together with training and maintenance

Connecting data

- Data sources within the enterprises are connected to the reasoning engine, via streaming data connectors. External web data can also be extracted automatically.
- For knowledge extraction from unstructured text, a combination of Machine Learning, NLP and reasoning can be leveraged to perform named-entity recognition, entity linking, schema and ontology mapping.

Deep Reason.ai Use Case 1

Central Bank Regulation Compliance/Fraud Detection

Snippet example of rule set

x **controls** y if
 x **directly holds** over 50% of y , or
 x **controls a set** S of companies that **jointly control** y

```
control(X,Y) :- own(X,Y,W), W > 0.5.  
control(X,Z) :- control(X,Y), own(Y,Z,W),  
                 V=msum(W,<Y>), V > 0.5.
```


DeepReason.ai Use Case 2

Knowledge Creation — Fact Enrichment and Constraints

Example Domain Constraint: Non-overlapping CEOs

Any company in the KG can only have a single CEO at any point in time (with possible missing data on start and end dates)

```
ceoPairOK(Organization, Job1, Job2) :-  
    ceoPositionWithEndDate(Organization, Job1, EndDate),  
    ceoPositionWithStartDate(Organization, Job2, StartDate),  
    EndDate < StartDate.
```

Core constraint

```
ceoPairOK(Organization, Job2, Job1) :-  
    ceoPairOK(Organization, Job1, Job2).
```

Inverse rule

```
kg:violation(Organization, Job1, Job2) :-  
    ceoPosition(Organization, Job1),  
    ceoPosition(Organization, Job2),  
    not ceoPairOK(Organization, Job2, Job1).
```

Deriving the
constraint violation

Thank you for your attention!



Questions