

# Monitoring Entities in an Uncertain World: Entity Resolution and Referential Integrity

Steven Minton, InferLink Corporation

Sofus Macskassy, Fetch Technologies

Peter LaMonica, Air Force Research Laboratories

Kane See, InferLink Corporation

Craig Knoblock, USC/Information Sciences Inst.

Greg Barish, Fetch Technologies

Matthew Michelson, Fetch Technologies

Ray Liuzzi, Raymond Technologies

# Many "Steve Minton"s

● Steve Minton, InferLink

● Steven Minton, Stanford University

● Stephen Minton, Brain Surgeon

● Steve Minton, Fetch Technologies

● Steven Minton, convicted felon

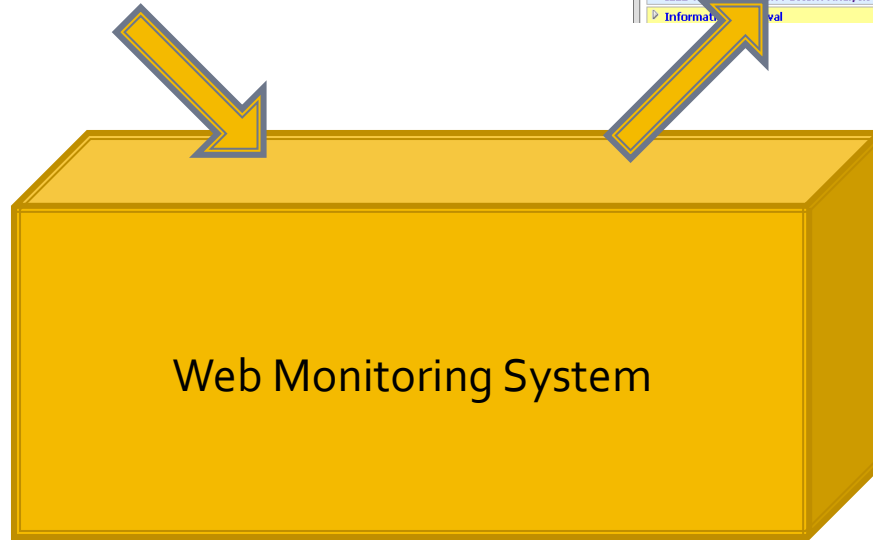
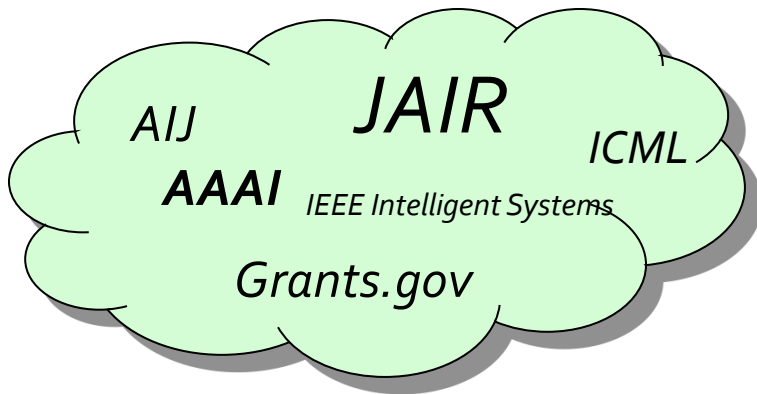
● Steven Minton, Jonosboro High School

● Steven Minton, JAIR

# Outline

- Application domain: Arms trafficking
- Entity Intelligence Portal (ENTEL)
- Entity resolution process
- Mistakes: Maintaining referential integrity

# Web Monitoring



**AI Portal** powered by Fetch Technologies

Welcome, editor! (sign out, my account) | Customize | Suggest a new source

Home    Publications    Funding    Administration

TIP: Give us your feedback on the portal

Publications

All    Full    Add Filter

[expand all]

Source	Last updated
ACM Transactions on Information Systems	Today
AI Communications	55 days ago
AI Magazine	124 days ago
Artificial Intelligence	17 days ago
Artificial Life	10 days ago
Cognitive Science	56 days ago
Computational Linguistics	18 days ago
Data Mining and Knowledge Discovery	48 days ago
Electronic Transactions on Artificial Intelligence	116 days ago
IEEE Intelligent Systems	11 days ago
IEEE Transactions on Knowledge and Data Engineering	12 days ago
IEEE Transactions on Pattern Analysis and Machine Intelligence	20 days ago
Information Systems	6 days ago

Ads by Google

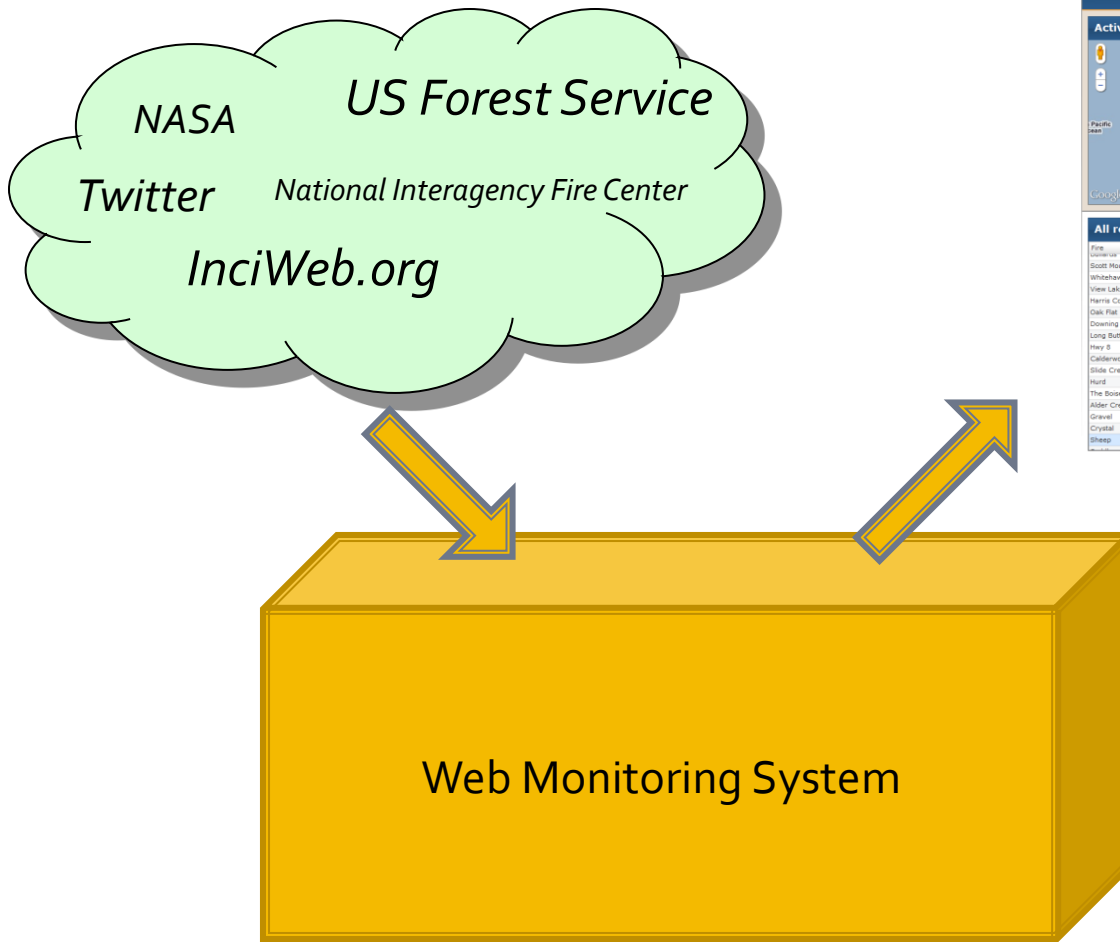
Wireless products, ZigBee Coordinator, Router & End devices Custom applications & profiles  
[www.bitech.se](http://www.bitech.se)

IEEE 1149.1/JTAG/BSca Intelligent solutions for extended JTAG testing and programming  
[www.goepet.com](http://www.goepet.com)

Online Risk Monitoring Security, Fraud Prevention, Channel Optimization and Brand Protection  
[www.eyelliance.com](http://www.eyelliance.com)



# Web Monitoring



**wildfire**  
Fetch TECHNOLOGIES

**Active**

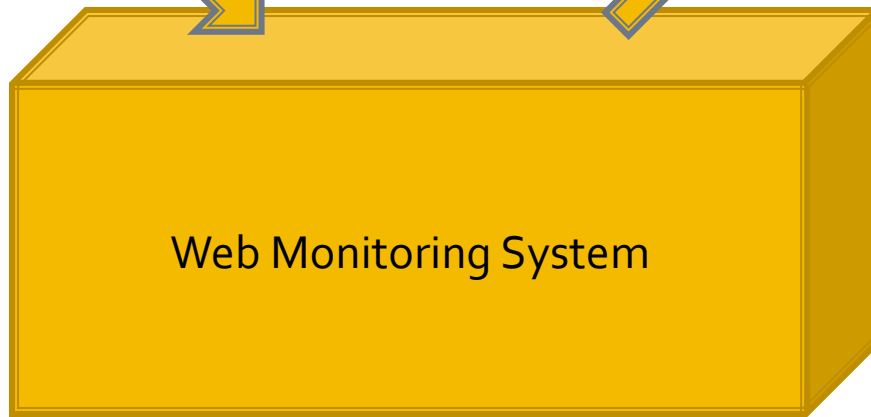
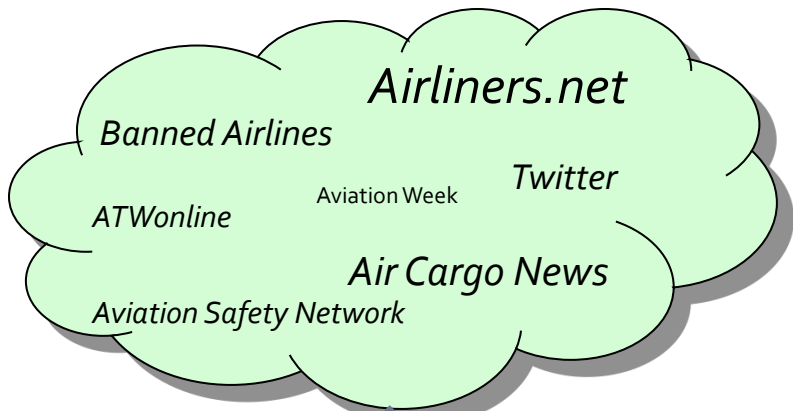
**Statistics**

- Active fires: 38
- Acres burning: 45,801
- Active acres: 88,694

**All recent**

Fire	Updated	Size	Status
Urbana	2010-09-01	1,269 acres	100% contained
Scott Mountain	2010-09-01	2,454 acres	30% contained
Whitehawk Complex	2010-09-01	1,439 acres	0% contained
View Lake Complex	2010-09-01	4,478 acres	35% contained
Harris Complex	2010-09-01	2,122 acres	20% contained
Dick Fire	2010-09-01	6,513 acres	70% contained
Downing Mountain	2010-09-01	315 acres	75% contained
Long Butte	2010-09-01	306,113 acres	100% contained
Hay 6	2010-09-01	2,019 acres	100% contained
Callowood	2010-09-01	255 acres	30% contained
Slide Creek	2010-09-01	989 acres	100% contained
Hurd	2010-09-01	1,331 acres	80% contained
The Boise Complex Update	2010-09-01	6 acres	0% contained
Alder Creek	2010-09-01	726 acres	100% contained
Gravel	2010-09-01	503 acres	0% contained
Crystal	2010-09-01	120 acres	0% contained
Sheep	2010-09-01	2,000 acres	0% contained

# Web Monitoring



### Fetch Intelligence Portal

by Fetch Technologies

home watchlists upcoming news known issues feedback about


Search: Search database...

#### DB Record for Plane [Cubana] spotted at Airport [Recife - Guararapes]

Original link: [www.airliners.net](http://www.airliners.net)  
*airlinersphoto*  
Published: 2010-12-08T23:18:02-08:00  
Source: [airlinersphoto](#)

Airline: [Cubana](#)  
Airport: [Recife - Guararapes](#)  
Aircraft: [Ilyushin Il-62M](#)  
Location: [Brazil](#)

CU-T1284 (cn 4053732) CU-T1284. Very rare plane! Talking off to Cuba in a cloudy morning. Beautiful plane...I just love it! Not the best of quality, but a beautiful and special photo to me. Check out the condensation over the wings. [Nikon D5000 - 70/300 f4.5.6 - ISO 1600]



Name	Predicate	Value
Cubana	Name	Cubana
Cubana	UsesAircraft	Ilyushin Il-62M
Cubana	Country	Brazil
Recife - Guararapes	IATA	REC
Recife - Guararapes	ICAO	SBRF
Recife - Guararapes	Name	Recife - Guararapes
Recife - Guararapes	Country	Brazil

# Victor Bout



# Arrested and Extradited in 2010



Charged with conspiracy to support a terrorist organization, money laundering, .....

# Some Companies Linked with Bout

Ishtar Airlines  
Irbis Air  
Omega Aircompany  
Mega Airlines  
WING AIR  
Aerocom  
Air Cess  
Galaxy Air  
Norse Air Charter  
Air Foyle  
Centrafricain Airlines  
Click Airways  
Air Bas  
Air Pass  
Anikay (Anikai) Airlines  
Pietersberg Aviation Services Systems  
Santa Cruz Imperial  
Great Lake Business Company  
Balkh Airlines  
Phoenix Aviation  
Air Zory  
JetLine International  
Dolphin Air  
MaxAvia  
Flying Dolphin  
San Air General Trading  
Sitrat Air  
Air Mero  
African Express  
Air Leone  
Inter Transavia

# One Thread: Air Cess to Air Bas

Registration	Aircraft Type	Construction Nbr	Previous Reg.	Sighting	Markings
UN-75002	Ilyushin 18E	185008603	3C-KKR	SHJ 11May03 SHJ 04Nov03	no markings a/w, n/t
UN-75003	Ilyushin 18V	184006903	3C-KKJ	SHJ 12Oct03	blue tail, no m/s
UN-75004	Ilyushin 18D	186009202	3C-KKK	SHJ 14Sep02 SHJ 04Nov03 SHJ 28Dec03	green cheatline and blue tail No t/t, blue tail all white
UN-75005	Ilyushin 18D	187010204	3C-KKL	SHJ 04Nov03	
UN-11007	Antonov 12B	9346509	3C-OOZ	SHJ Oct02 SHJ 11May03 DXB 12Oct03	No m/s all white c/s no titles

*[From Ruudleeuw.com]*



# An Online Data Source

Photos: Ilyushin Il-18D Aircraft Pictures | Airlines.net - Windows Internet Explorer

http://www.airliners.net/photo/Yuzhnaya-Aircompany/Ilyushin-Il-18D/0635957/M/

6 of 11




Photo Copyright © Graham Penson AIRLINERS.NET

6 of 11

Rate and comment on this photo by clicking on the star rating below:

★★★★★ ★★★★★ ★★★★★ ★★★★★ ★★★★★

Aircraft	Taken at
<a href="#">More: Yuzhnaya Aircompany</a> <a href="#">More: Ilyushin Il-18D</a>	<a href="#">More: Karachi - Jinnah International (Quaid-e-Azam) (KHI / OPKC)</a> <a href="#">More: Pakistan, July 27, 2004</a>
Remark	Photographer
UN-75001 (cn 187009904) Possibly operated as a pax/cargo combi, this IL18D was being loaded on the Karachi ramp.	<a href="#">More: Graham Penson</a> <a href="#">Contact Graham Penson</a> <a href="#">↔ Link to me!</a>

mind + body + spirit

WE'VE GOT PLENTY OF HANDS TO TAKE CARE OF THEM.

REPLAY

# Another Source

Antonovs AN-8 and AN-12 at Sharjah - Windows Internet Explorer

http://www.ruudleeuw.com/an8\_and\_12.htm

★ Favorites Antonovs AN-8 and AN-12 at Sharjah



This AN-12B 3C-OOZ (c/n 9346509) still wears the colourful scheme of Ural Airlines, but it seems likely that it is now operated by someone else. I took this photo april 26th, 2000 at Sharjah and as it was still around in 2001, one can assume that Sharjah has become its homebase. They probably have a reason for not advertising the name of the company on the aircraft, pity.....





# One More Source

ASN Aircraft accident Antonov 12B UN-11007 Mukalla-Al-Rayyan International Airport (MKX) - Windows Internet Explorer

http://www.aviation-safety.net/database/record.php?id=20050331-2

AviationSafetyNetwork FLIGHT SAFETY FOUNDATION  
an exclusive service of Flight Safety Foundation www.flightsafety.org

Home » Database » 2005

### Accident description

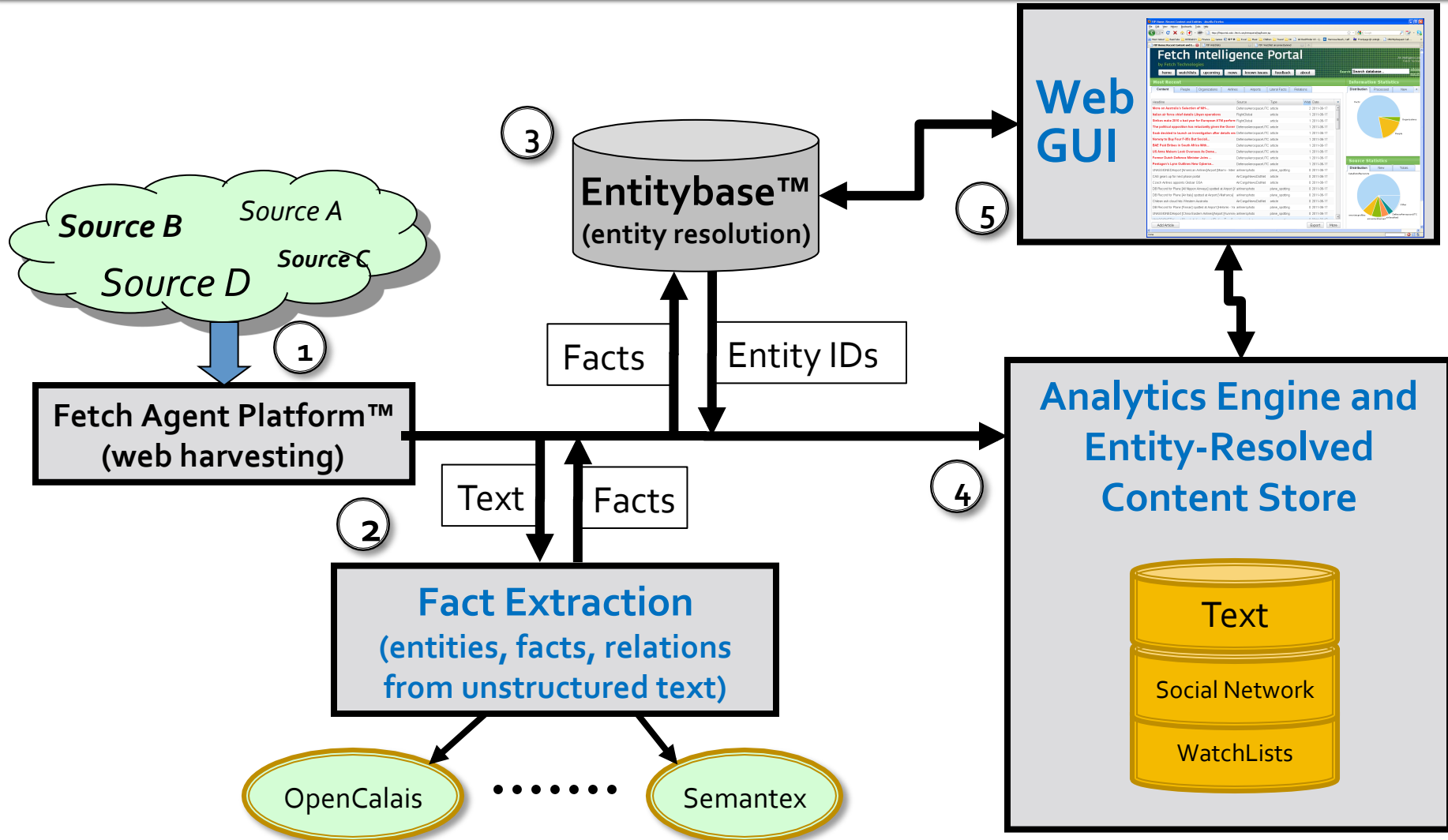
languages:

[Share](#) [Tweet](#) 0 [+7](#) 0

**Status:**  
**Date:** 31 MAR 2005  
**Type:** [Antonov 12B](#)  
**Operator:** [RPS Air Freight](#)  
**Registration:** UN-11007  
**C/n / msn:** 9346509  
**First flight:** 1969  
**Engines:** 4 Ivchenko AI-20M  
**Crew:** Fatalities: 0 / Occupants: 8  
**Passengers:** Fatalities: 0 / Occupants: 0  
**Total:** Fatalities: 0 / Occupants: 8  
**Airplane damage:** Written off  
**Airplane fate:** Written off (damaged beyond repair)  
**Location:** Mukalla-Al-Rayyan International Airport (MKX) ([Yemen](#))  
**Phase:** Takeoff (TOF)  
**Nature:** Cargo  
**Departure airport:** [Mukalla Airport \(MKX\) \(MKX\)](#), Yemen  
**Destination airport:** [Dubai Airport \(DXB\) \(DXB/OMDB\)](#), United Arab Emirates

**Narrative:**  
The Antonov had made a refueling stop at Mukalla and was bound for Dubai, carrying 7 tons of fish. The takeoff from Mukalla had to be aborted, but the plane could not be stopped on the runway. It overran by 400 meters and caught fire. The fire was quickly contained. Possible hull-loss.

# ENTEL Architecture



# Fetch Intelligence Portal

by Fetch Technologies

An Intelligence portal by Fetch Technologies

- home
- watchlists
- upcoming
- news
- known issues
- feedback
- about

Search:  advanced search

## Most Recent

## Information Statistics

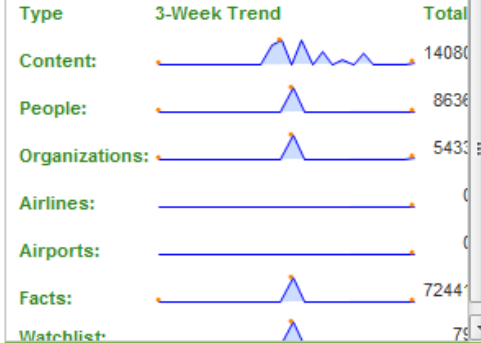
- Content
- People
- Organizations
- Airlines
- Airports
- Literal Facts
- Relations

- Distribution
- Processed
- New

Headline	Source	Type	Watr	Date
A330/340 change to inhibit autopilot if airspeed unreliable	FlightGlobalNews	article	0	2011-08-05
Bombardier nets Gulf Air order	FlightGlobalNews	article	0	2011-08-05
Alitalia narrows half-year loss to €94 million	atwonline	article	0	2011-08-05
Frontier reaches tentative accord with flight attendants	atwonline	article	0	2011-08-05
Congress adjourns with FAA partial shutdown unresolved	atwonline	article	0	2011-08-05
Atlas Air Worldwide's first-half profit dips 48% but ACMI re...	atwonline	article	0	2011-08-05
Dempsey confirmed as next top US mili...	DefenseAerospaceUTC	article	0	2011-08-05
Serbia seeking investors for new airline to replace Jat	atwonline	article	0	2011-08-05
Investigators probe Caribbean Airlines 737NG runway over...	atwonline	article	0	2011-08-05
Obama picks new deputy defense chief...	DefenseAerospaceUTC	article	0	2011-08-05
Debt deal has little effect on Pentag...	DefenseAerospaceUTC	article	0	2011-08-05
French investigators: Air France 447 pilots didn't fully grasp...	atwonline	article	0	2011-08-05
AA and BAA parent could merge	AirCargoNewsDotNet	article	0	2011-08-05
Perceptive MPs have concluded that reducing defense spe...	DefenseAerospaceUTC	article	0	2011-08-05
Lawmakers Look To Contain Defense Cuts	AviationWeek	article	0	2011-08-05
Yusen to appeal cartel fine	AirCargoNewsDotNet	article	0	2011-08-05
Singapore Engine Plant Starting With Trent 900s	AviationWeek	article	0	2011-08-05
Rio-Paris jet crash probe 'discredite...	DefenseAerospaceUTC	article	0	2011-08-05

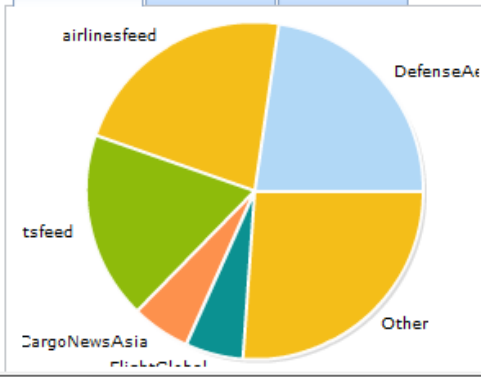
Add Article

Export More



## Source Statistics

- Distribution
- New
- Totals



# Fetch Intelligence Portal

by Fetch Technologies

home

watchlists

upcoming

news

known issues

feedback

about

Search: S

## Watchlist detail for aircarriersbanned

Matches

Members

Entity Name	Content Description	Content Type	Content Source	Date Publish
Volga-Dnepr Airlines	Will Cargo Profiling Take Hold?	article	AirCargoNewsCom	2011-07-27
Volga-Dnepr Airlines	Big Lift To Subcontinent	article	AirCargoNewsCom	2011-07-27
Wings Air	Lion Air, Virgin Blue Add 33 ATR-72 ...	article	DefenseAerospac...	2011-07-27
SonAir	Heli-Expo: Sikorsky Sells 12 Helicop...	article	DefenseAerospac...	2011-07-27
SonAir	Cargo Handlers Get No Respect	article	AirCargoNewsDot...	2011-07-27
SonAir	Prosperous Bosphorous IATA Cargo Week	article	AirCargoNewsDot...	2011-07-27
Volga-Dnepr Airlines	Russian investigators have determined that the Tupolev...	article	FlightGlobal	2011-07-27
Philippine Airlines	Philippine carriers and Iran Air feature on latest EU blac...	article	FlightGlobal	2011-07-27
Air Koryo	Philippine carriers and Iran Air feature on latest EU blac...	article	FlightGlobal	2011-07-27
Ctilink	Air Canada, Garuda, Thai and Air Austral could subleas...	article	FlightGlobal	2011-07-27
Volga-Dnepr Airlines	Hearts & Flowers In SMSS SFO	article	AirCargoNewsDot...	2011-07-27
Kartika Airlines	Farnborough International Airshow News	article	atwonline	2011-07-27
Kartika Airlines	Kartika Airlines Orders 30 Sukhoi Su...	article	DefenseAerospac...	2011-07-27
Ctilink	Airbus Wins 58 New Orders, Options a...	article	DefenseAerospac...	2011-07-27
Philippine Airlines	AirAsia to launch Philippines-based affiliate	article	atwonline	2011-07-27
Mandala Airlines	Indonesia's Mandala Airlines suspends operations, see...	article	atwonline	2011-07-27

# Fetch Intelligence Portal

by Fetch Technologies

home

watchlists

upcoming

news

known issues

feedback

about

Search: Search database...

## DB Record for Plane [Cubana] spotted at Airport [Recife - Guararapes]

Original link: [www.airliners.net](http://www.airliners.net)

*airlinersphoto*

Published: 2010-12-08T23:18:02-08:00

Source: [airlinersphoto](#)

Airline: [Cubana](#)

Airport: [Recife - Guararapes](#)

Aircraft: Ilyushin Il-62M

Location: Brazil

CU-T1284 (cn 4053732) CU-T1284. Very rare plane! Takking off to Cuba in a cloudy morning. Beautiful plane...i just love it! Not the best of quality, but a beautiful and special photo to me. Check out the condensation over the wings. [Nikon D5000 - 70/300 f/4:5.6 -

Remark: ISO 1600]



People

Organizations

Literal Facts

Relations

Name	Predicate	Value
Cubana	Name	Cubana
Cubana	UsesAircraft	Ilyushin Il-62M
Cubana	Country	Brazil
Recife - Guararapes	IATA	REC
Recife - Guararapes	ICAO	SBRF
Recife - Guararapes	Name	Recife - Guararapes
Recife - Guararapes	Country	Brazil

## PT Merpati Nusantara Airlines(AIRLINE)

First seen: 2011-07-08T12:35:33-07:00

Last seen: 2009-07-06T07:00:00-07:00

Times seen: 5

Split Entity

Content

Literal Facts

Relations

Add Facts

Watchlists

Aliases

Airports

Routes

Headline ^	Source	Type	Watchlist	Date	
DB record for Merpati Nusantara Airlines	aircarriersbanned	structured_input	0	2011-07-08	
DB record for Merpati Nusantara Airlines	airlinesfeed	structured_input	0	2011-07-08	
<b>DB Record for Plane [Merpati Nusantara Airlines] spo...</b>	airlinersphoto	plane_spotting	1	2011-07-08	
<b>Merpati abandons purchase deal with Xi'an aircraft</b>	CargoNewsAsia	article	2	2009-07-06	



## PT Merpati Nusantara Airlines(AIRLINE)

First seen: 2011-07-08T12:35:33-07:00

Last seen: 2009-07-06T07:00:00-07:00

Times seen: 5

Split Entity

Content

**Literal Facts**

Relations

Add Facts

Watchlists

Aliases

Airports

Routes

Name ▼	Predicate	Value	Last Seen	Frequency	▼
PT Merpati Nusantara Airlines	Name	Merpati Nusantara Airlines	2011-07-08	2	▲
PT Merpati Nusantara Airlines	UsesAircraft	Xian MA-60	2011-07-08	1	
PT Merpati Nusantara Airlines	Country	China	2011-07-08	1	
PT Merpati Nusantara Airlines	Callsign	MERPATI	2011-07-08	2	
PT Merpati Nusantara Airlines	IATA	MZ	2011-07-08	2	
PT Merpati Nusantara Airlines	ICAO	MNA	2011-07-08	2	≡
PT Merpati Nusantara Airlines	Country	Indonesia	2011-07-08	2	
PT Merpati Nusantara Airlines	Carrier	Merpati Nusantara Airlines	2011-07-08	1	
PT Merpati Nusantara Airlines	Name	PT Merpati Nusantara Airlines	2009-07-06	2	
PT Merpati Nusantara Airlines	Nationality	Indonesian	2009-07-06	1	
PT Merpati Nusantara Airlines	Name	Indonesian state-owned airline...	2009-07-06	1	
PT Mernati Nusantara Airlines	Name	Mernati	2009-07-06	2	▼

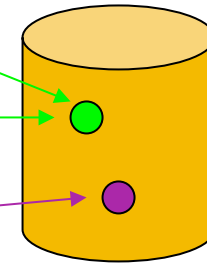
# The Entity Resolution Problem

- Entity resolution: Link incoming records describing the same entity from multiple sources

*R. Landis, President, Fetch Technologies*

*Robert Landes, CEO, Fetch Software*

*R. Land, CEO, French Alliance Technologies*



- Many “common sense” issues, for instance:
  - Multiple formats for names, addresses, etc.
    - R.L. Landes vs. Robert Landes
  - Noisy, incorrect values
    - Landes vs. Landis
  - Multi-valued attributes
    - Landes can be both President and CEO
  - Aliases and Deception



# Fetch Intelligence Portal

by Fetch Technologies

An intelligence portal by  
Fetch Technologies

home

watchlists

upcoming

news

known issues

feedback

about

Search:

advanced  
search

## Qatar's Al Baker hits back at AEA's unfair competition claim

**Original link:** <http://www.flightglobal.com/articles/2011/02/10/353051/qatars-al-baker-hits-back-at-aear-unfair-competition.html>

**nws\_article**

**Published:** 2011-02-10T22:01:08-08:00

**Source:** FlightGlobal

Outspoken Qatar Airways boss Akbar Al Baker has responded angrily to a recent address by Association of European Airlines secretary general Ulrich Schulte-Strathaus, in which he called help from ICAO in tackling the growing threat from the Gulf carriers.

Outspoken Qatar Airways boss Akbar Al Baker has responded angrily to a recent address by Association of European Airlines secretary general Ulrich Schulte-Strathaus, in which he called help from ICAO in tackling the growing threat from the Gulf carriers. He also pointed out that with these carriers integrated vertically into commerce, tourism and foreign policy, their governments treat the airlines as "just part of that chain".

Schulte-Strathaus added that this policy is incongruent with those of several regions and countries, including the USA, the European Union, Australia and Japan. He suggested that new global aviation policies were needed and that ICAO should be responsible for developing that framework. Stating that the AEA boss's comments are "factually incorrect and unfounded", Al Baker says he has "no problem" for ICAO to create a multilateral system similar to how the World Trade Organisation tackles free trade. "I would go even further: Why won't we apply the WTO principles on aviation to ensure that first and foremost the rights of the consumers and free competition apply across the board?" Al Baker points out that most of the countries listed by Schulte-Strathaus are signatories of the Agenda for Freedom brokered by IATA, which calls for free market access. "Does he advocate that this policy of liberalisation should only apply when his member airlines are the beneficiaries and not vice-versa?" he says.

To compare the Gulf carriers' huge order book to the US airlines' fleets is "extremely erroneous", says Al Baker as domestic operations are the backbone of the US air transport industry. "The portion of the capacity deployed internationally as a ratio of the total capacity deployed by the US airlines is minimal," he says.

Al Baker also challenges the AEA's view about the Gulf airlines being part of a vertically integrated economic chain: "Can [Schulte-Strathaus] tell me of any country which does not consider its air transport industry, be that an individual airline or a multitude of them, as part of national interests?" The Qatar Airways chief points to the financial assistance US airlines received from their government after 9/11 "in order to ensure the continuity of the US air transport industry" as well as "the billions of Euros that are allowed under the EU laws to pump into airlines for restructuring purposes".

People Organizations **Literal Facts** Relations

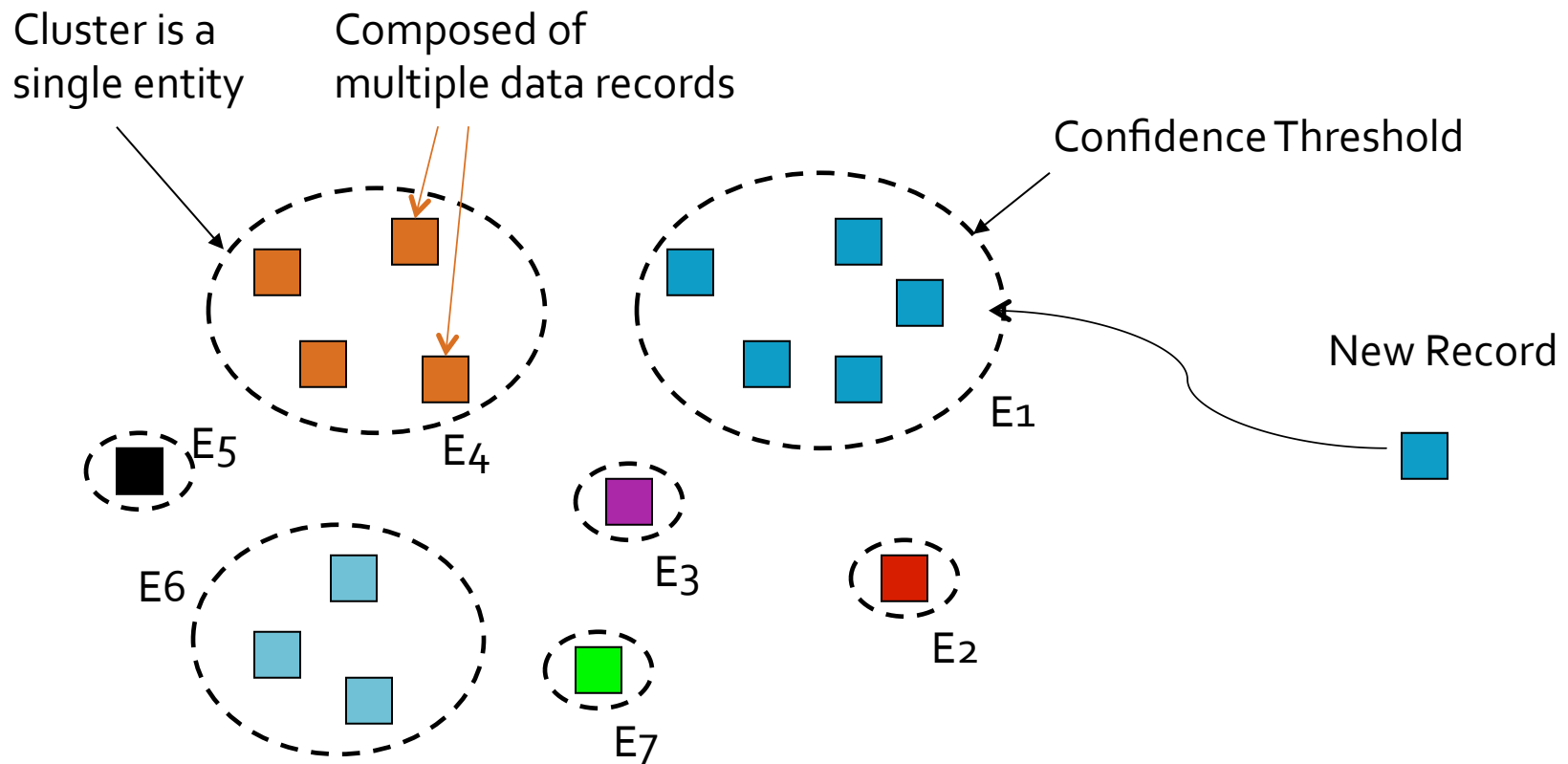
Name	Predicate	Value	Last Se
Emirates	Ticker	EMIRA	2011-02
Ethihad Airways	Name	Ethihad Airways	2011-02
Ethihad Airways	Name	Ethihad Airways P.J.S.C.	2011-02
EU	Name	European Union	2011-02
EU	Type	governmental civilian	2011-02
EU	Name	EU	2011-02
European Airlines	Name	European Airlines	2011-02
Qatar Airways	Name	Qatar Airways	2011-02
Qatar Airways	Nationality	Qatari	2011-02
Qatar Airways	Nationality	Qatari	2011-02
Qatar Airways	Name	Qatar Airways W.L.L.	2011-02
Qatar Airways	Name	<u>Outspoken Qatar Airways</u>	2011-02
Schulte-Strathaus	Name	Ulrich Schulte-Strathaus	2011-02
Schulte-Strathaus	CareerField	political	2011-02
Schulte-Strathaus	Name	Schulte-Strathaus	2011-02
Schulte-Strathaus	Name	M. Schulte-Strathaus	2011-02

Export

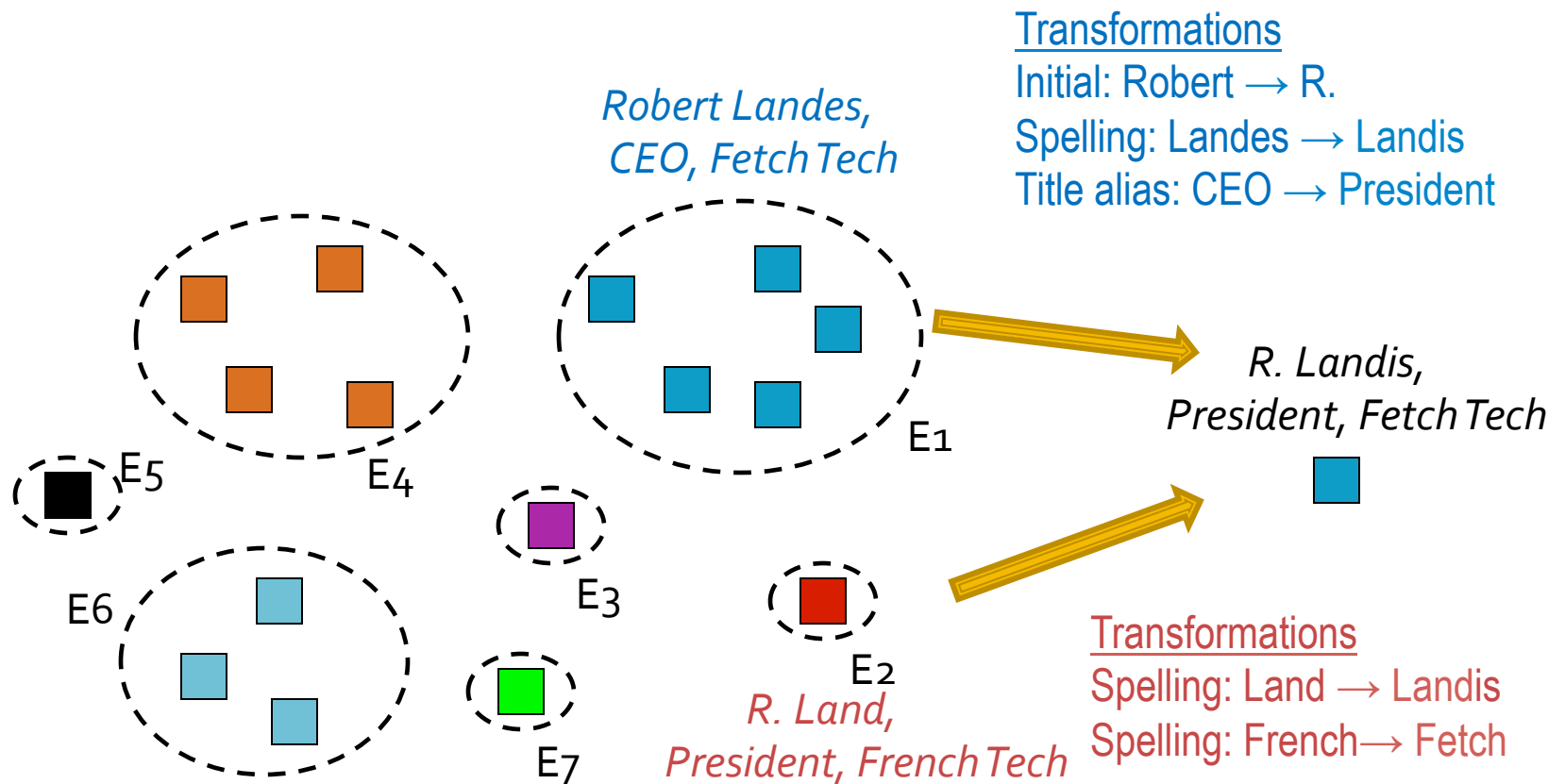
More

Outspoken Qatar Airways

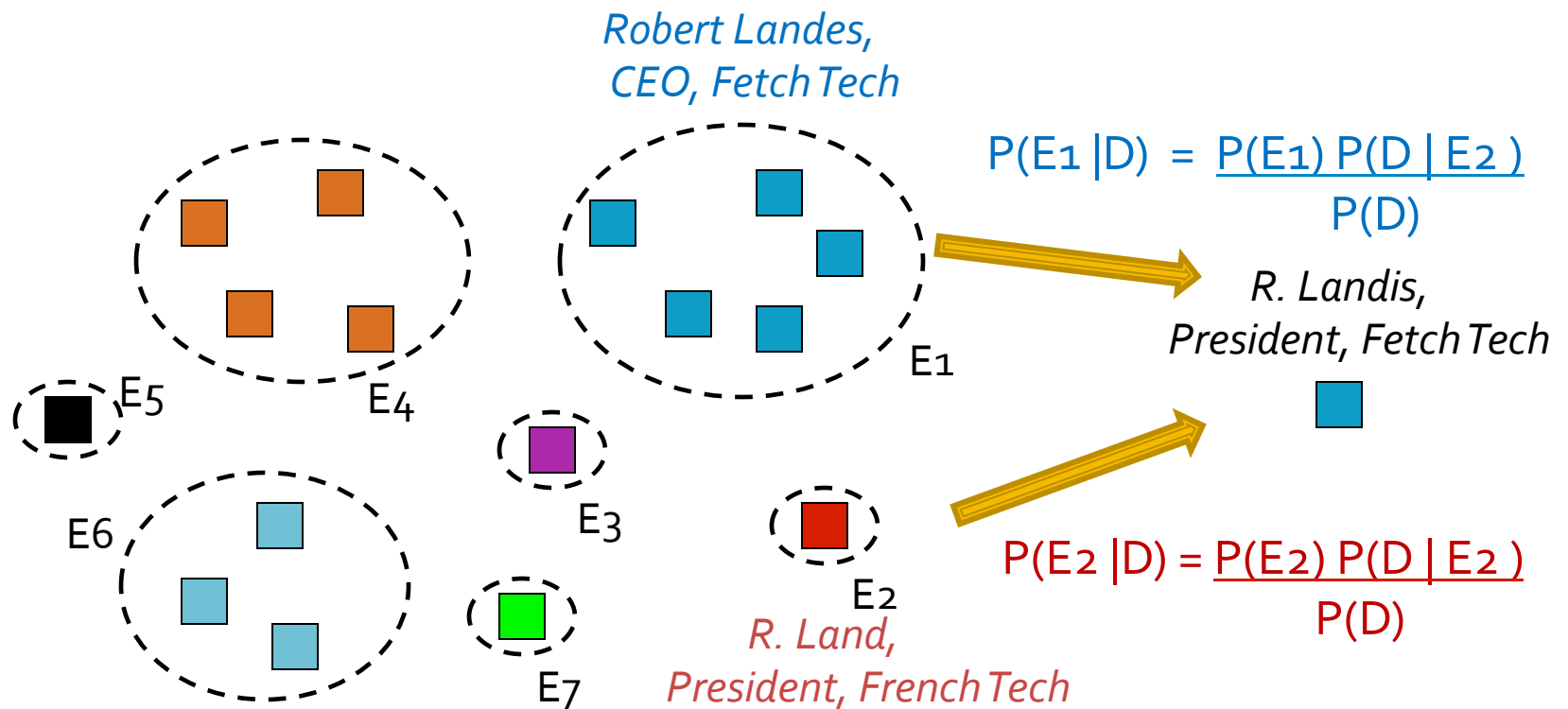
# Clustering in EntityBase



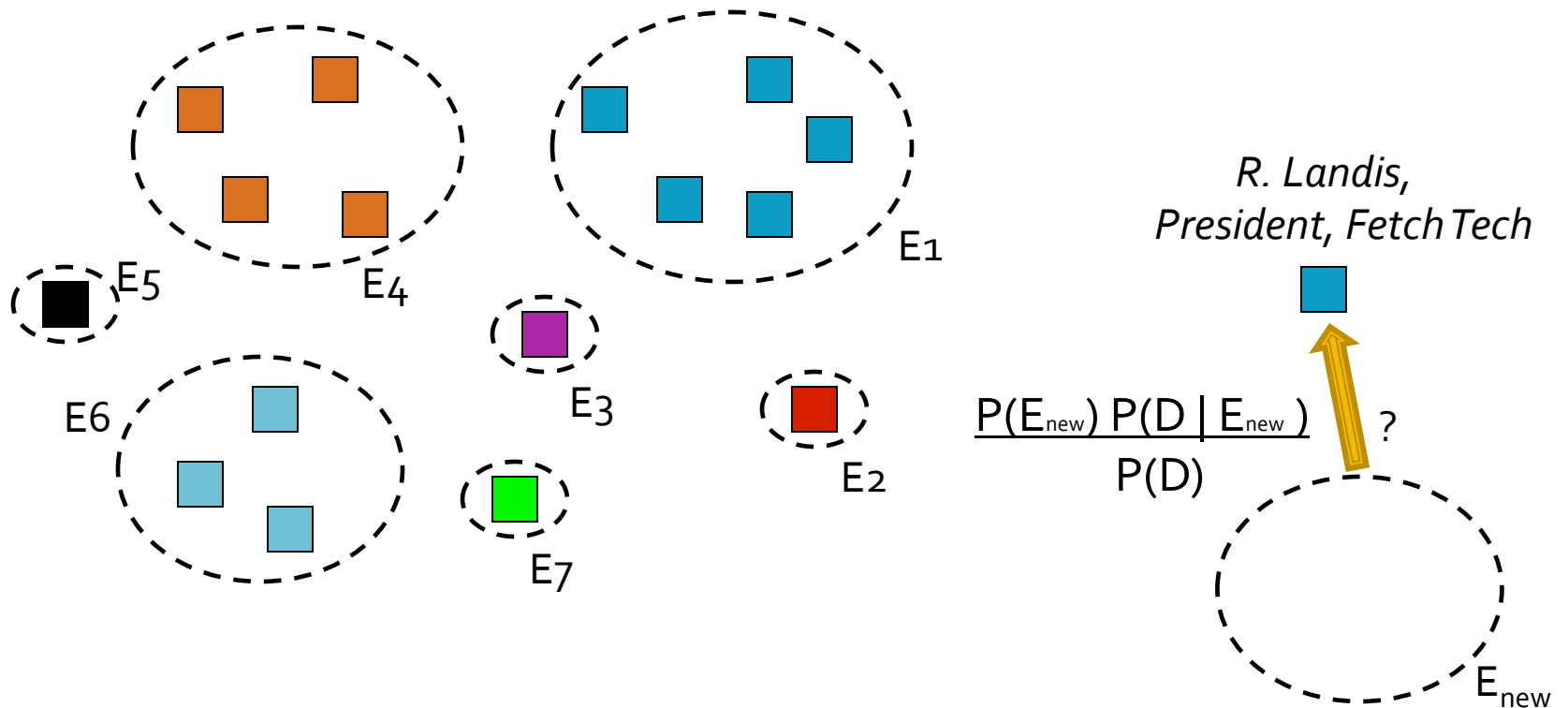
# Estimating Probabilities



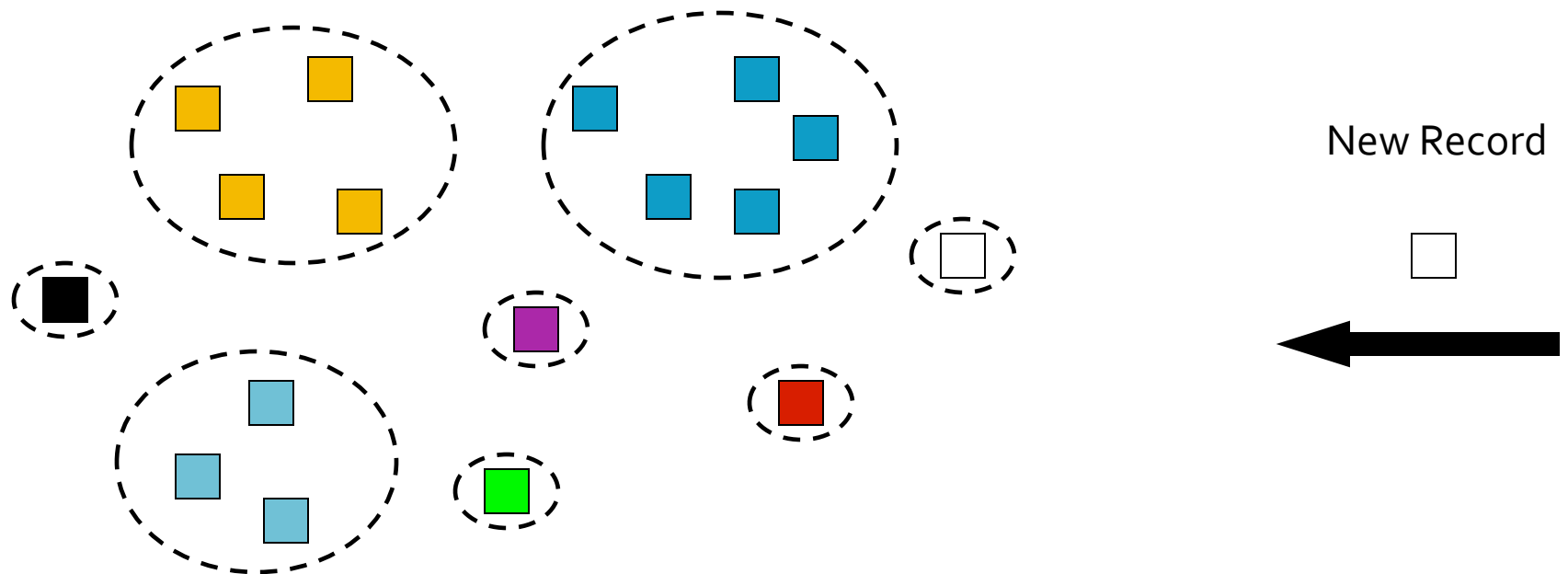
# Estimating Probabilities



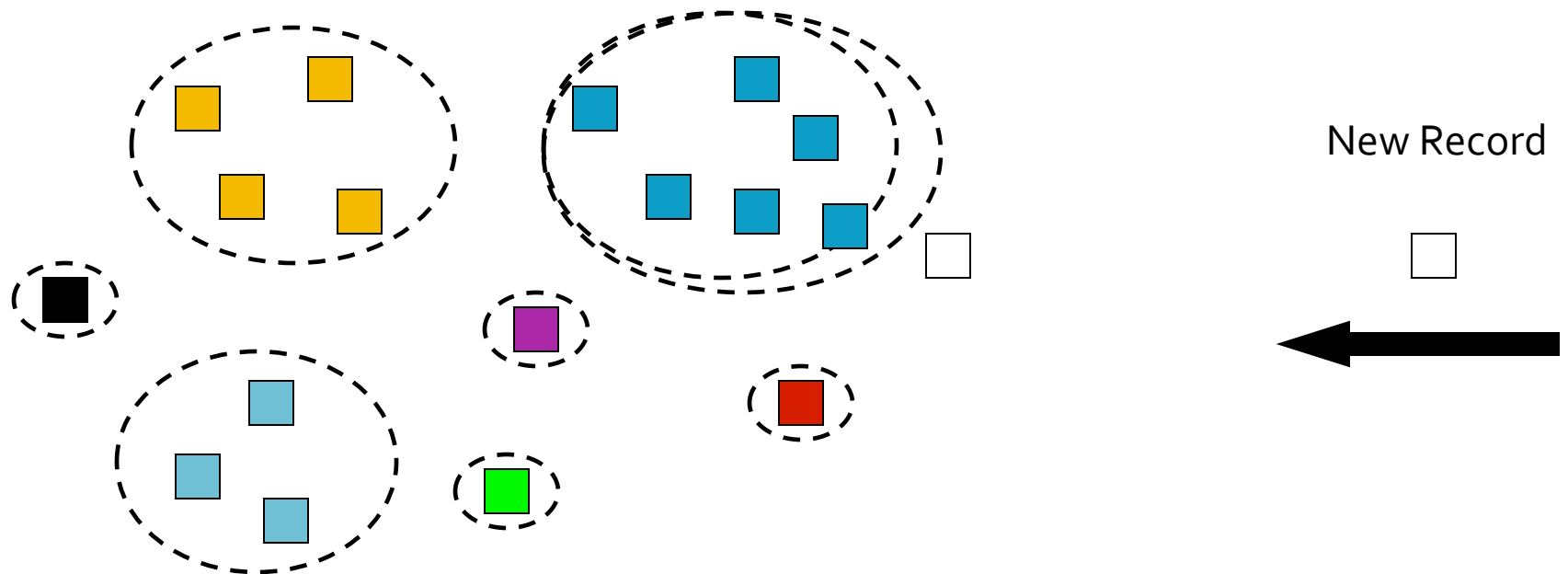
# Estimating Probabilities



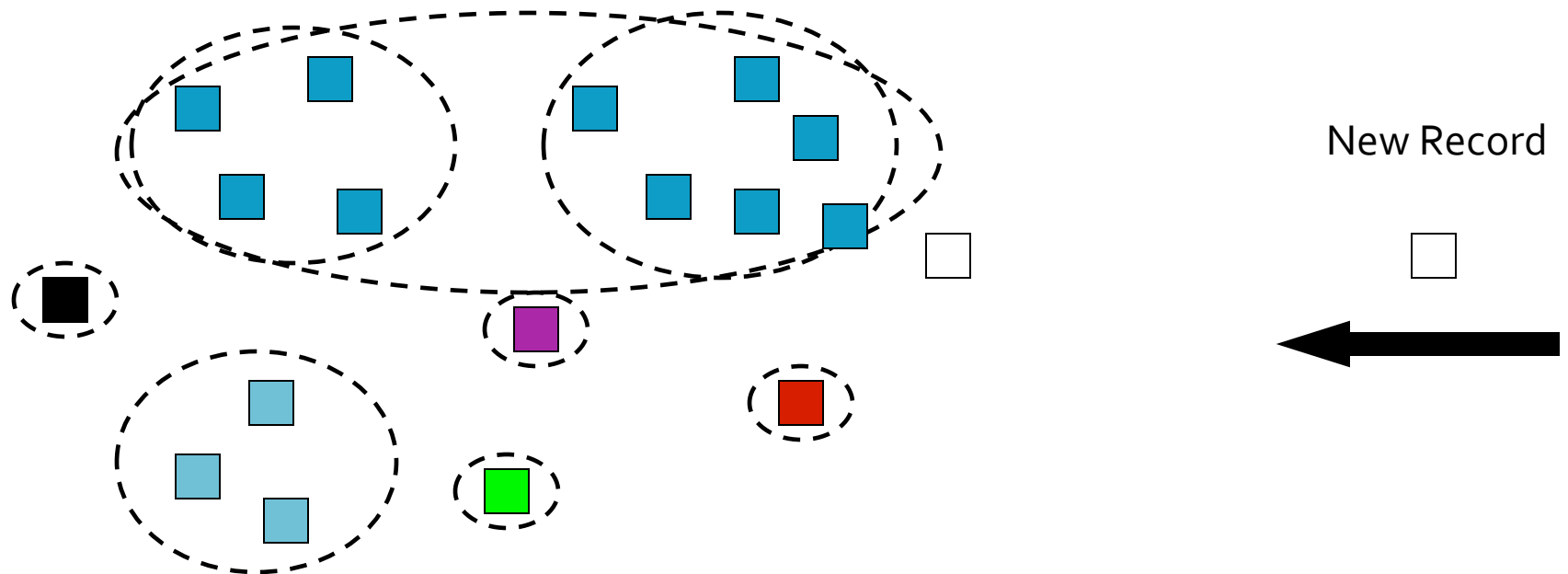
# Clustering



# Clustering



# Clustering

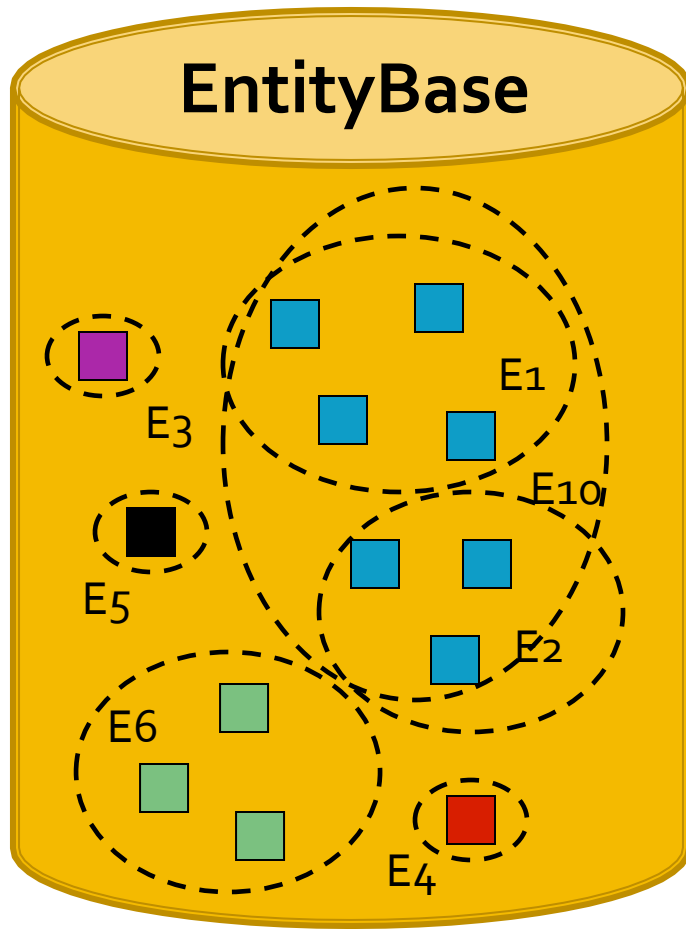




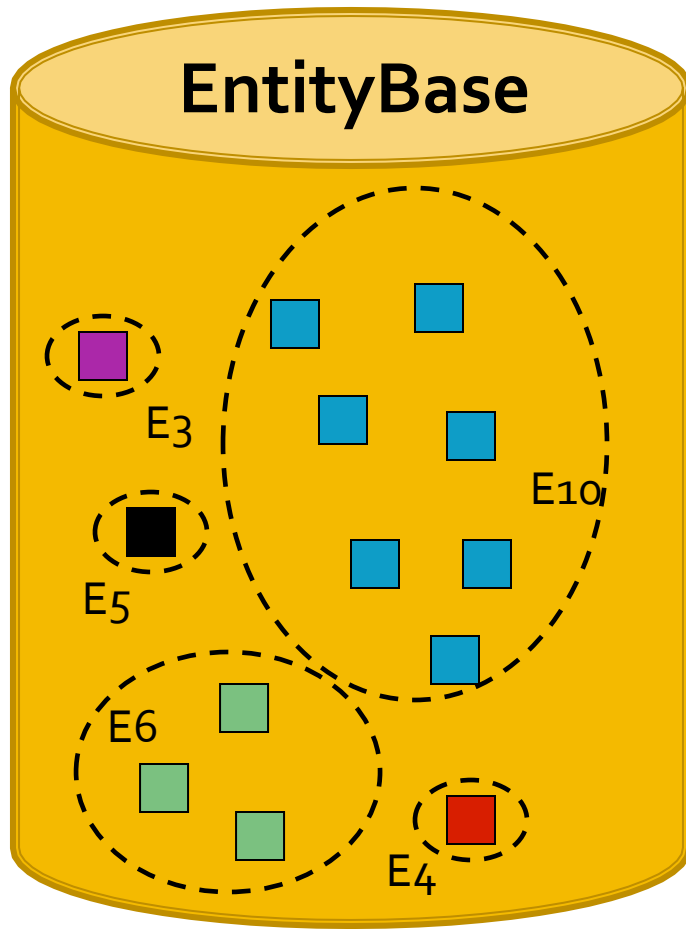
# Correcting mistakes : Merging and splitting clusters

- Merge example:
  - Air Cess and Air Bas aircraft
- Split example:
  - George H. W. Bush and George W. Bush

# Merging Two Clusters



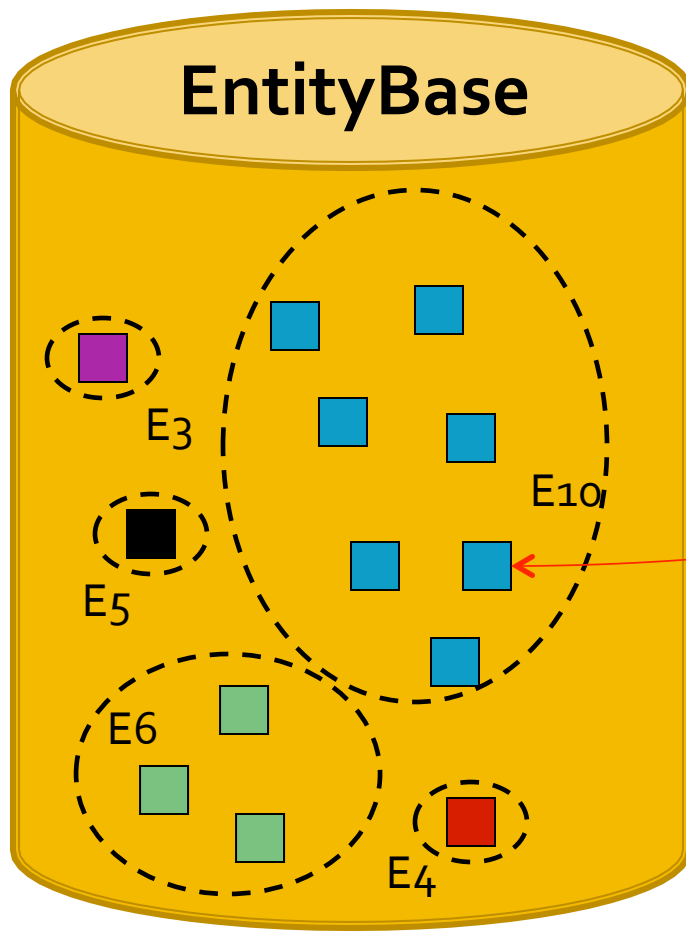
# Referential Integrity



The screenshot shows a web browser window displaying the Fetch Intelligence Portal for PT Merpati Nusantara Airlines. The page includes navigation links (home, watchlists, upcoming, news, known issues, feedback, about) and a table of facts. The table has columns for Name, Predicate, Value, Last Seen, and Frequency. The value "Kurniyo - Muliyo" is highlighted in red, and "E2 ?" is written below it.

Name	Predicate	Value	Last Seen	Frequency
PT Merpati Nusantara Airlines	ServesAirport	Kurniyo - Muliyo	2011-07-08	1
PT Merpati Nusantara Airlines	Employee	Bambang Bhakti	2009-07-06	1

# Refer-by-Description



Fetch Intelligence Portal  
by Fetch Technologies  
home watchlists upcoming news known issues feedback about

PT Merpati Nusantara Airlines(AIRLINE)

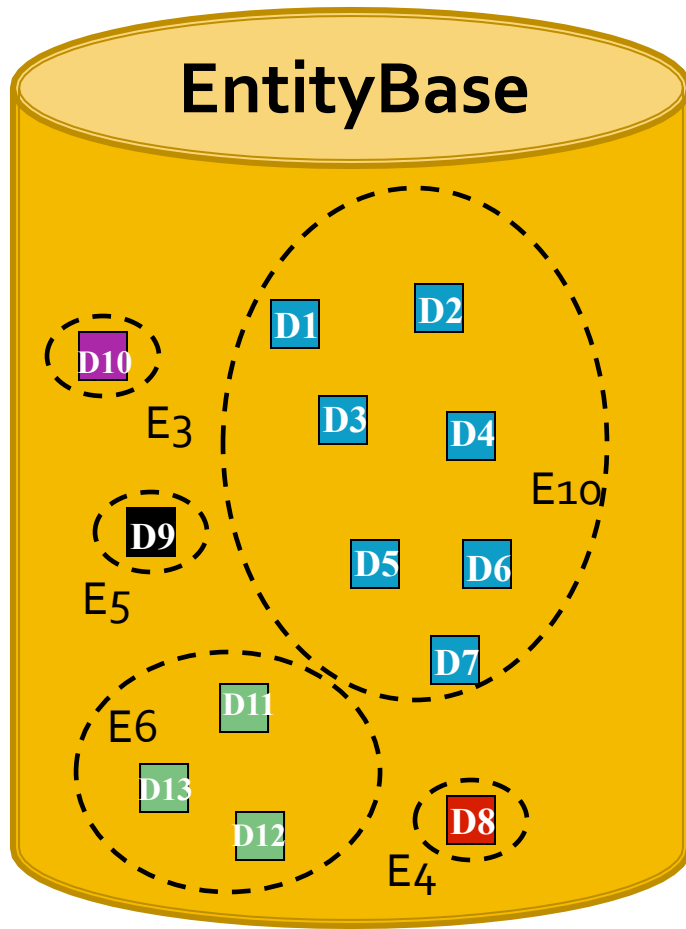
First seen: 2011-07-08T12:35:33-07:00  
Last seen: 2009-07-06T07:00:00-07:00  
Times seen: 5

Split Entity

Content Literal Facts Relations Add Facts Watchlists Aliases Airports Routes

Name	Predicate	Value	Last Seen	Frequency
PT Merpati Nusantara Airlines	ServesAirport	Kuning-Mulaha	2011-07-08	1
PT Merpati Nusantara Airlines	Employee	Banbang Bhakti	2009-07-06	1

# Refer-by-Description

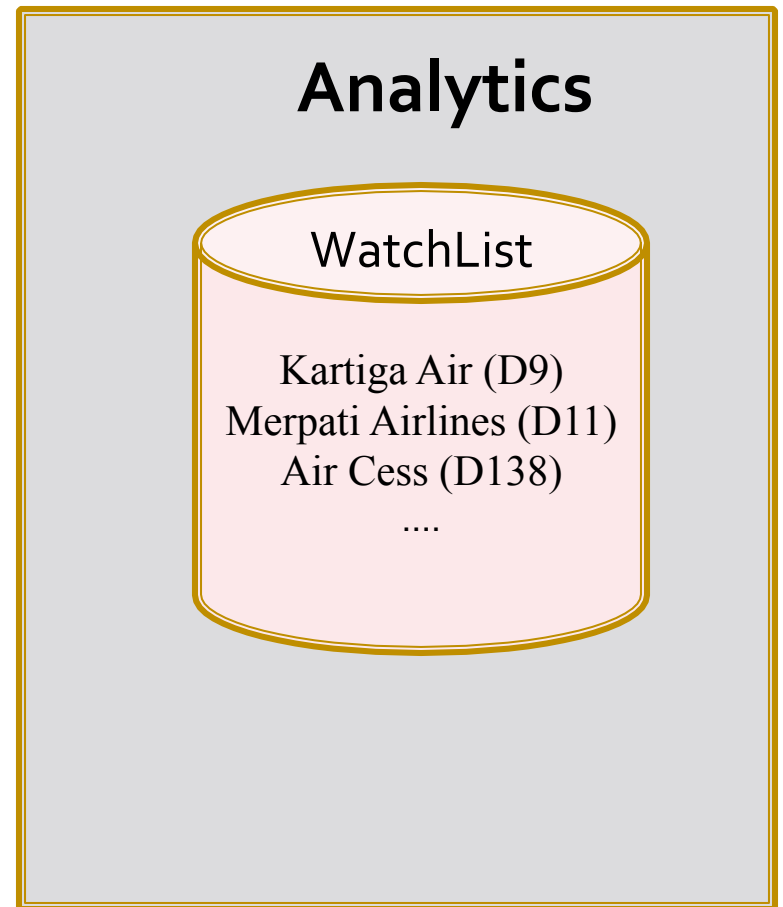
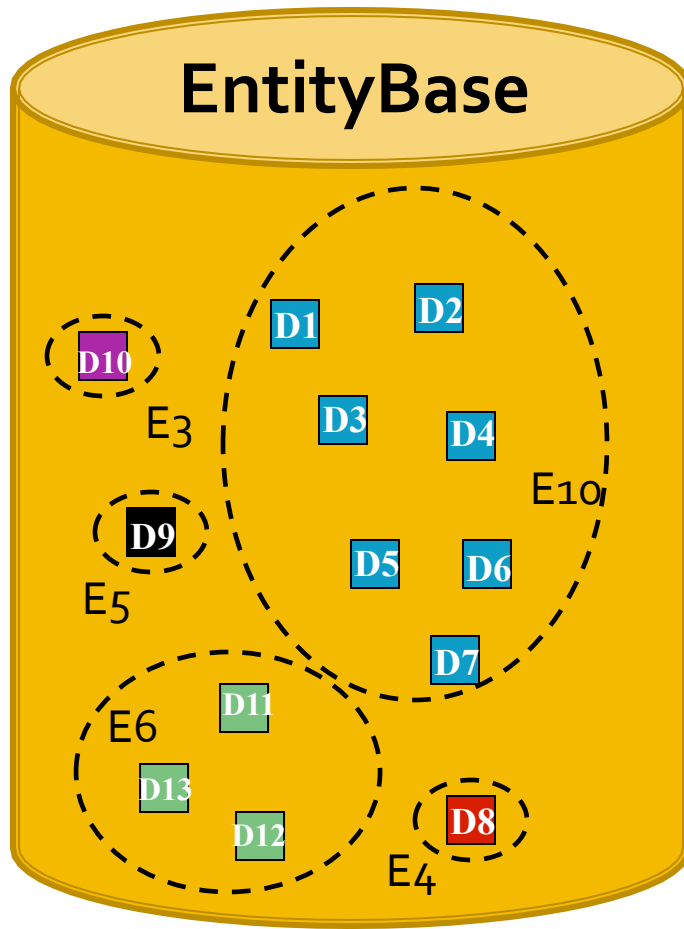


The screenshot shows a web browser window titled 'Fetch Intelligence Portal' for 'PT Merpati Nusantara Airlines (AIRLINE)'. The page includes navigation links like 'home', 'watchlists', 'upcoming', 'news', 'known issues', 'feedback', and 'about'. A table displays search results for the airline, with the value 'Bangbang Bhakti' highlighted in red. Below the table, the text 'D6' is displayed.

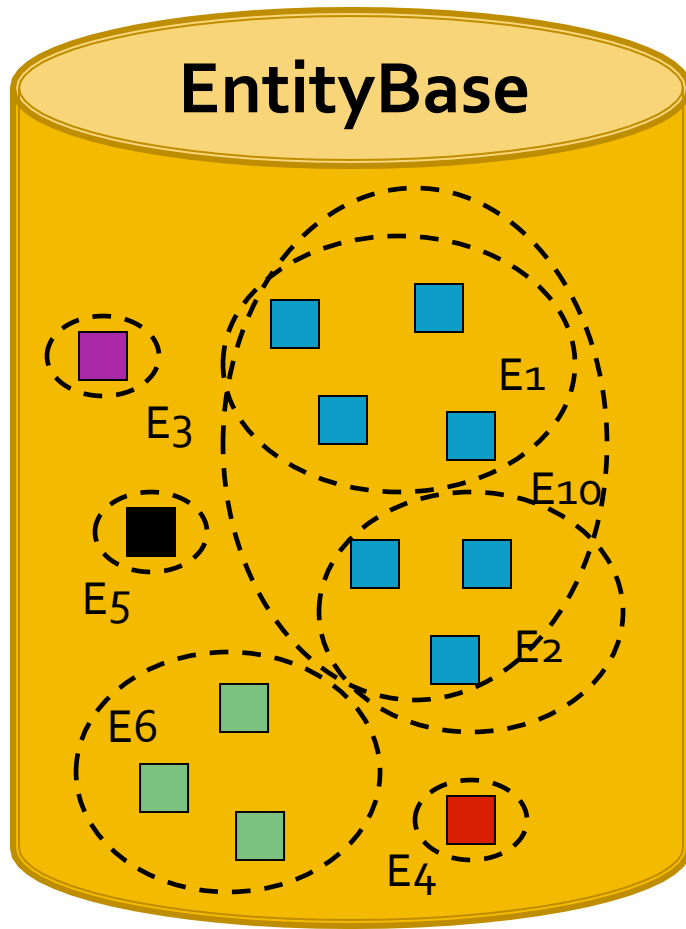
Name	Predicate	Value	Last Seen	Frequency
PT Merpati Nusantara Airlines	ServesAirport	Kuning-Majaba	2011-07-08	1
PT Merpati Nusantara Airlines	Employee	Bangbang Bhakti	2009-07-06	1

**D6**

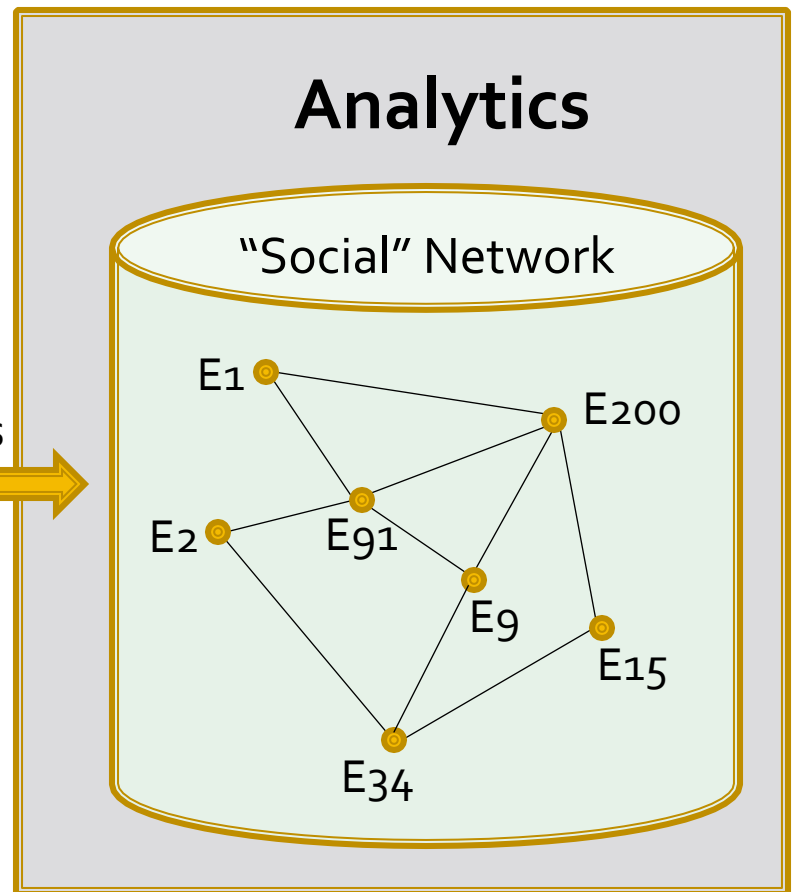
# Refer-by-Description



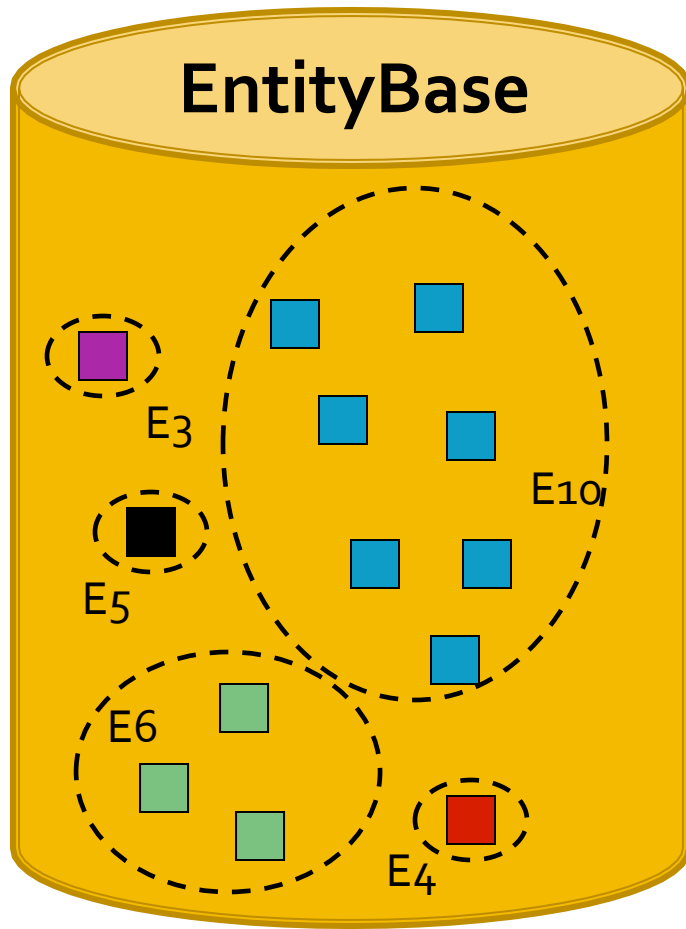
# Refer-by-Identifier



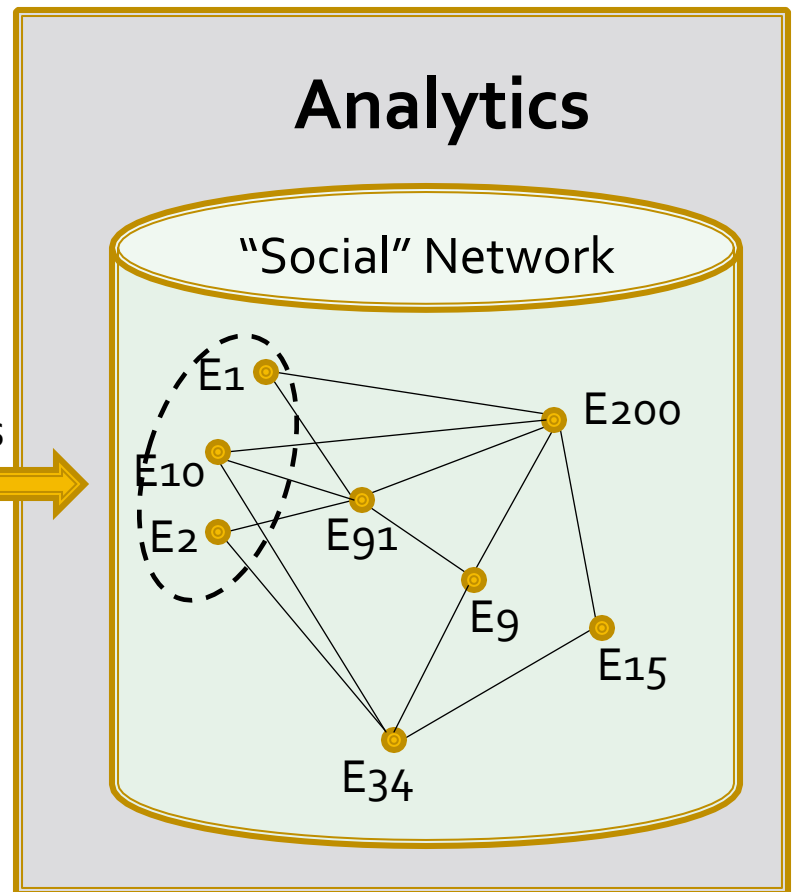
Publish  
Merges/Splits



# Refer-by-Identifier



Publish  
Merges/Splits





# Maintaining Referential Integrity

- Two approaches:
  - Refer-by-Description
    - Indirect reference: Point to a cluster member
    - Advantage: Easy, no synchronization necessary
    - ...But limits information that client can cache
  - Refer-by-Identifier
    - Direct reference: Cluster ID
    - Advantage: Client can cache arbitrary information
    - ...But client must synch with EntityBase and maintain consistency

# Why It Matters



- Vision: Entity Resolution in a decentralized world
  - E.g., the Semantic Web (Glaser, Jaffri & Millard, 2009)

# Summary

- Entity resolution can be hard: “AI Complete”
  - Arms trafficking domain
- Entity merges and splits will occur
- Entity resolution clients must be designed to deal with this
- Two strategies: Refer-by-Description and Refer-by-Identifier
- System status: Being evaluated by AF personnel