



SPARK
SUMMIT
2017

LEVERAGING APACHE SPARK TO DISRUPT AIRLINE PRICING DISTRIBUTION

EVERYMUNDO

EVERYMUNDO

WHO WE ARE AND WHAT WE DO

INTRO TO OUR DATA ECOSYSTEM

2 SPECIFIC EXAMPLES OF SPARK USAGE



© 2017 EveryMundo

EVERYMUNDO

EVERYMUNDO

BASED IN MIAMI, FL

10 YEARS OLD

6 YEARS WORKING WITH AIRLINES

23 AIRLINE CUSTOMERS TO DATE



© 2017 EveryMundo

EVERYMUNDO



WE ARE IMMIGRANTS - 65 % BORN OUTSIDE THE US
57 MUNDOS
11 NATIONALITIES

Datum / Date: 6.7.17

OUR WHY -
TRANSFORM HOW COMPANIES INNOVATE
THROUGH OUR TECHNOLOGY & KNOWLEDGE

OUR WHY -
TRANSFORM HOW COMPANIES INNOVATE
THROUGH OUR TECHNOLOGY & KNOWLEDGE

WHAT WE DO WITH AIRLINES...

TRANSFORM HOW AIRLINES COMMUNICATE
WITH CUSTOMERS ONLINE

OUR WHY -
TRANSFORM HOW COMPANIES INNOVATE
THROUGH OUR TECHNOLOGY & KNOWLEDGE

WHAT WE DO WITH AIRLINES...

TRANSFORM HOW AIRLINES COMMUNICATE
WITH CUSTOMERS ONLINE

Spark IS ESSENTIAL

INFORM THROUGH ECOMMERCE PAGES

The screenshot shows a flight search interface on the airberlin.com website. At the top, there's a navigation bar with links like "Home", "Flights", "Hotels", "Cars", "Travel Services", "Airlines", and "Log in". Below the navigation, there are search filters: "One Way" and " +/- 3 Days", followed by a link to "Advanced Flight Search". The main search form includes fields for "From" (Miami FL), "To" (Berlin - Tegel), "Departure" (05/08/2017), "Return" (05/15/2017), and "Passengers" (1 Passenger). A red "Search" button is on the right.

Top 3 c Low Fare Calendar*

*PRICES INC *PRICES INCLUDE ROUND-TRIP FLIGHTS.

MAY

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
30	1	2	3	4	5	6
		USD 1031 Return 05/09	USD 1704 Return 05/10		USD 1579 Return 05/10	USD 1534 Return 05/24
7	8	9	10	11	12	13
USD 1343 Return 05/23	USD 1316 Return 05/11	USD 982 Return 05/19				
14	15	16	17	18	19	20
	USD 1489 Return 05/25		USD 1059 Return 05/30		USD 1068 Return 05/29	USD 1142 Return 06/11
21	22	23	24	25	26	27
	USD 929 Return 05/30	USD 1122 Return 06/08		USD 1270 Return 05/29	USD 978 Return 05/29	
28	29	30	31	1	2	3
USD 928 Return 05/04	USD 895 Return 05/07					

Datum / Date:

6.7.17

INFORM THROUGH ECOMMERCE PAGES

The screenshot shows the Ryanair website interface for searching flights to the Greek Islands. At the top, there's a search bar with fields for 'From' (Departure Airport), 'To' (Destination Airport), 'By Date' (DEPARTURE), 'By Airline' (RYANAIR), 'Passenger' (1 Passenger), and a 'Let's go!' button. Below this, a large banner says 'Flights to the Greek Islands'. A section titled 'Choose deals for flights to Greek Islands' lists several flight options:

Flight Details	Price	One way	
Glasgow GLA - Chania CHQ Fly out Wed 17 May	£ 39.99	Chania (CHQ) 7 Jun	£ 41.99
Birmingham BHX - Chania CHQ Fly out Fri 19 May	£ 24.99	Physical Tue 27 Jun	£ 41.99
London STN - Chania CHQ Fly out Sat 10 Jun	£ 69.99	London STN - Chania CHQ Fly out Sun 11 Jun	£ 35.99
London STN - Corfu CFU Fly out Mon 13 Jun	£ 51.99	London STN - Corfu CFU Fly out Sun 17 Jun	£ 48.99
Birmingham BHX - Corfu CFU	£ 19.99	Gatwick LGW - Corfu CFU Physical Sun 27 May	£ 32.99
London STN - Corfu CFU Fly out Thu 29 Jun	£ 51.99	Greece ATH - Corfu CFU ATH	£ 21.99
London STN - Kefalonia EFL Fly out Thu 29 Jun	£ 63.99	Greece ATH - Kefalonia EFL ATH	£ 35.99
Kat. Malaga LPA - Rhodes RHO Fly out Mon 29 Jun	£ 41.99	London STN - Rhodes RHO Physical Thu 29 Jun	£ 63.99
Palma PMI - Chania CHQ Fly out Wed 21 May	£ 50.99	Athens ATH - Rhodes RHO ATH	£ 19.99
Athens ATH - Mykonos JMK Fly out Mon 19 Jun	€ 19.99		

Datum / Date:

6.7.17

COMMUNICATE THROUGH SEARCH

A screenshot of a search results page from a search engine. The search bar at the top contains the query "flights to berlin". Below the search bar are several search results:

- airberlin.com - Fly from Miami to Berlin - Nonstop Flights from \$ 379**
Ad www.airberlin.com/Flights/Berlin ▾
4.5 ★★★★★ rating for airberlin.com
We offer our guests the highest level of comfort, service, and entertainment.
- Fly to Berlin - Fly Lufthansa - lufthansa.com**
www.lufthansa.com/Berlin ▾
Fly Lufthansa from Miami to Berlin. Great offers available.
- Flights from Miami, FL (MIA) to Berlin, Germany (all airports)** Sponsored ▾
www.google.com/flights
Miami, FL (MIA) Berlin, Germany (all airports)
Tue, May 16 Tue, May 20
Nonstop airberlin 8h 26m from \$1,025
Connecting Multiple airlines 11h 55m+ from \$1,040
United 11h 20m+ from \$1,055
Other airlines
- Flights from Miami to Berlin from USD 758 | airberlin**
https://flights.airberlin.com ▾ ... ▾ Flights to Germany ▾ Miami - Berlin ▾
Flying from Miami. Miami International Airport (MIA) is a major hub between the United States and Europe. Book your flights from Miami through airberlin and enjoy a relaxing journey to your destination. The airport is not far from the city – about 14 kilometers.

Customer: “I’m looking for a flight to Berlin.”

Airline: “I see you’re in Miami. We offer flights from Miami to Berlin. Here is our best flight now.”

APPEAL THROUGH DISPLAY

The Telegraph

HOME | NEWS | SPORT | BUSINESS | ALL SECTIONS 

Travel | Destinations

Destinations | Hotels | Offers | Holiday types | City | Beach | Tours | Cruise | Ski | Family | Luxury | Advice

Travel Destinations Europe Greece Articles

Greek Islands travel guide



A holiday in Greece conjures up images of idyllic islands and long sandy beaches

By **Marc Dubin**, DESTINATION EXPERT

WINTER GETAWAYS!
FROM **€19.99**



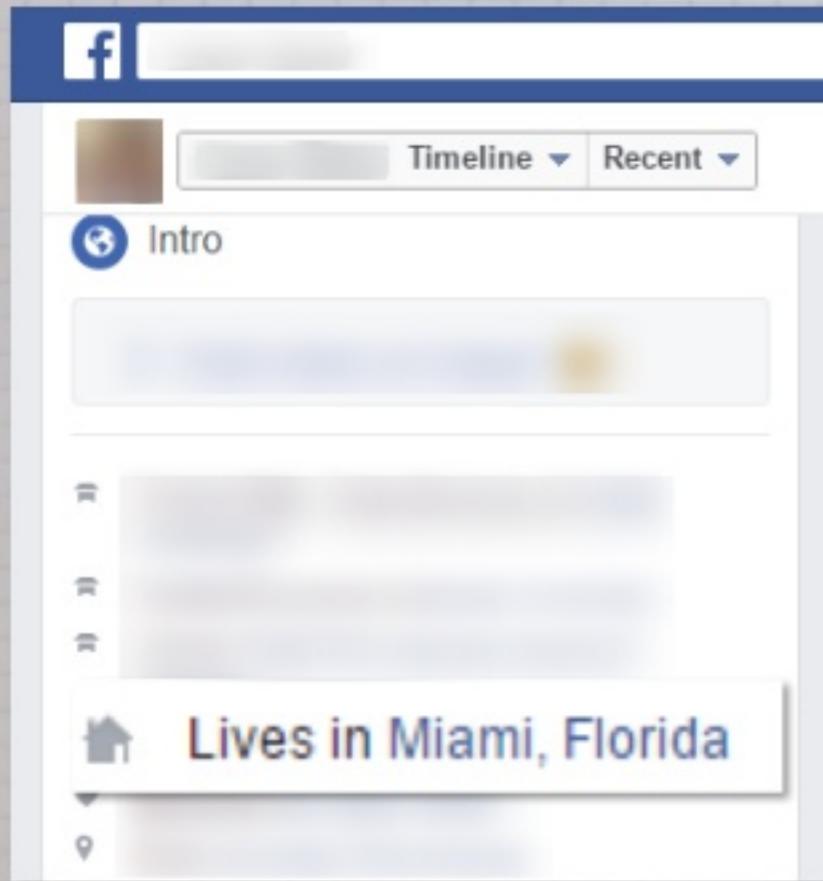
BOOK NOW! ▶

 **RYANAIR**
LOW FARES. MADE SIMPLE.
For full terms and conditions visit [ryanair.com](#)

Datum / Date:

6-7-17

PAY ATTENTION IN SOCIAL



A screenshot of a Facebook search results page for the query 'machu picchu'. The search bar at the top shows the query. Below it is a navigation bar with tabs for 'Top', 'Posts', 'People', 'Photos', 'Videos', 'Pages', 'Places', 'Groups', and 'Apps'. On the left, there's a 'POSTED BY' filter with options like 'Anyone' (selected), 'You', 'Your Friends', 'Your Groups', 'Your Friends and Groups', and 'Choose a Source...'. To the right, there's a 'Pages' section with a result for 'Machu Picchu - Peru'. It shows a thumbnail image of Machu Picchu, the page's name, its category as 'Local/Travel Website', the number of likes (159,242), and a link to the page's profile.

Datum / Date:

6.7.17

PAY ATTENTION IN SOCIAL



LATAM AIRLINES

Destinations & Offers Plan & Book Check-In & Other Services Travel Information Frequent Flyer

Home / Flights / United States - Peru

Flights from United States to Peru **from \$499***

One Way Round-trip Flying to more than

From City or Airport To City or Airport Depart 04/20/2017 Return 04/24/2017 GUEST 1 Passenger Show Flights

Best fares on flights from United States to Peru
BEST FARES FOR ROUND-TRIP FLIGHTS FOUND BY OTHERS IN THE LAST 48 HOURS

From Miami
Depart 05/09/2017, Return 05/28/2017

From	To	From	To
From Miami	To Lima (LIM)	From \$499*	To Lima (LIM)
From Miami	To Bogota (BOG)	From \$784*	To Bogota (BOG)
From Miami	To Quito (QUI)	From \$655*	To Quito (QUI)
From Miami	To La Paz (LPZ)	From \$994*	To La Paz (LPZ)
From Washington, D.C.	To Miami	From \$609*	To Miami
From Washington, D.C.	To Bogota (BOG)	From \$646*	To Bogota (BOG)
From Los Angeles	To Lima (LIM)	From \$917*	To Lima (LIM)
From New York	To Bogota (BOG)	From \$985*	To Bogota (BOG)
From New York	To Lima (LIM)	From \$878*	To Lima (LIM)

Datum / Date: 6.7.17

KEEP IN TOUCH WITH EMAIL

FRONTIER
LOW FARES DONE RIGHT

BOOK A FLIGHT MANAGE TRAVEL VIEW IN BROWSER

**GET OUT OF THE COLD
WITH BEARLY LOW FARES!
FROM DENVER**

\$ 34
one way

BOOK NOW

TRAVEL VALID TUESDAYS & WEDNESDAYS
Buy by Dec. 9, 2016. Fly by Mar. 8, 2017.
*Fare shown is one way for domestic, nonstop travel on specific days in select markets.
7 day advance purchase required. Seats are limited. Restrictions apply.

Datum / Date: 6-7-17

KEEP IN TOUCH WITH EMAIL

FRONTIER
LOW FARES DONE RIGHT

PLAN & BOOK MANAGE TRAVEL WAYS TO SAVE TRAVEL INFORMATION

FLIGHTS FROM DENVER (DEN) STARTING AT \$34*

From: To: Depart: Return: 1 Traveler



Home / Flights / From Denver

 **BEST AIRFARE DEALS FROM DENVER (DEN)**
*Prices found by others for one-way flights

Denver (DEN) To San Diego (SAN)
Departing May 23, 2017

**from
\$34***

Datum / Date: 6-7-17

INSTANT INTERACTIONS WITH DATA

United | United.com

Reservations Travel Information Deal & Offers MilesPlat®

Profile • Change • Status • My United

Deals on flights from Boston to Houston

PLACES	DATES	PRICES
BOS - IAH	03/27 - 04/01	Under \$900

Under \$900 Clear

From	Dates	Prices
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$1
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$186*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$159*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$146*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$188*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$188*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$146*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$188*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$159*
Boston (BOS) - Houston (IAH) Round-trip Economy	Depart Mar 27, 2017, Return Apr 1, 2017	from \$146*

*Prices have been available for round trips within the last 48 hours and may not be currently available. Additional baggage fees and fees for other optional service may apply. Other terms

Deals on flights from Boston to Houston

PLACES	DATES	PRICES
BOS - IAH	03/27 - 04/01	Under \$900

Under \$900 Clear

Datum / Date: 6-7-17

SPEAK ALL LANGUAGES & CURRENCIES

CATHAY PACIFIC

Discover Where We Fly Manage Latest Offers Travel Information Marco Polo Club Cathay Dragon

Find and book your next trip

Return One Way

Fuzhou, Fuzhou Changli Int'l (FOC) China	Hong Kong, Hong Kong Int'l (HKG) Hong Kong
DEPART ON On 05 May 2017	RETURN ON On 12 May 2017
PASSENGER CLASS 1 Passenger, Economy	Search Flights >

Flights to Hong Kong (HKG) from CNY600

Home > Flights > To Hong Kong

Hong Kong airfare

Book based on flight round-trip or return in the last 48 hours. (Please note that some fares may only apply when there are multiple travellers.)

Shanghai (SHA) to Hong Kong (HKG) Depart on 20/06/2017, Return on 25/06/2017	from CNY770
Hangzhou (HGH) to Hong Kong (HKG) Depart on 18/06/2017, Return on 21/06/2017	from CNY1,020
Nanjing (NKG) to Hong Kong (HKG) Depart on 18/06/2017, Return on 21/06/2017	from CNY950
Shanghai (PVG) to Hong Kong (HKG) Depart on 11/06/2017, Return on 14/06/2017	from CNY770
Beijing (PEK) to Hong Kong (HKG) Depart on 25/06/2017, Return on 05/07/2017	from CNY830
Chengdu (CTU) to Hong Kong (HKG) Depart on 22/06/2017, Return on 27/06/2017	from CNY970

CATHAY PACIFIC

简体中文

发现 | 香港机场 | 香港 | 马来西亚 | 新加坡 | 马可孛罗会 | 圆滑游轮

查找和预订您下一次的旅程

往返 单程

始发地: 杭州 杭州长乐国际机场 (FOC) 中国	目的地: 香港, 香港国际机场 (HKG) 香港
出发日期: 2017年05月05日 周五	返回日期: 2017年05月12日 周五
舱位: 头等舱 1 舱位, 经济舱	搜索航班 >

飞往香港 (HKG) 的航班由CNY 600起

主页 > 航班 > 香港

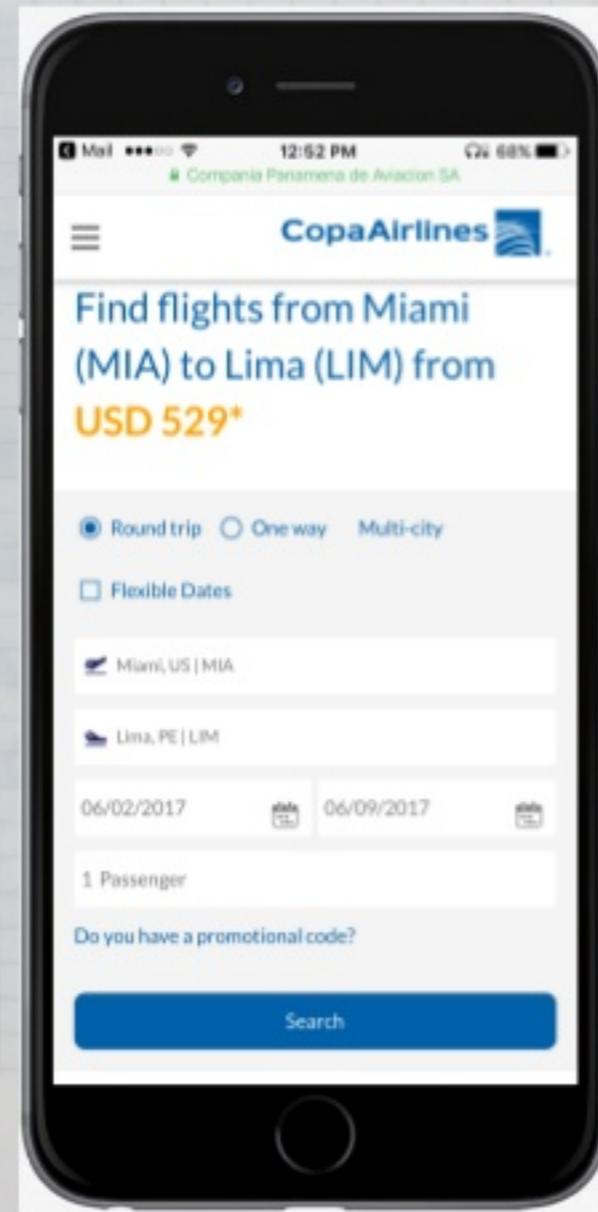
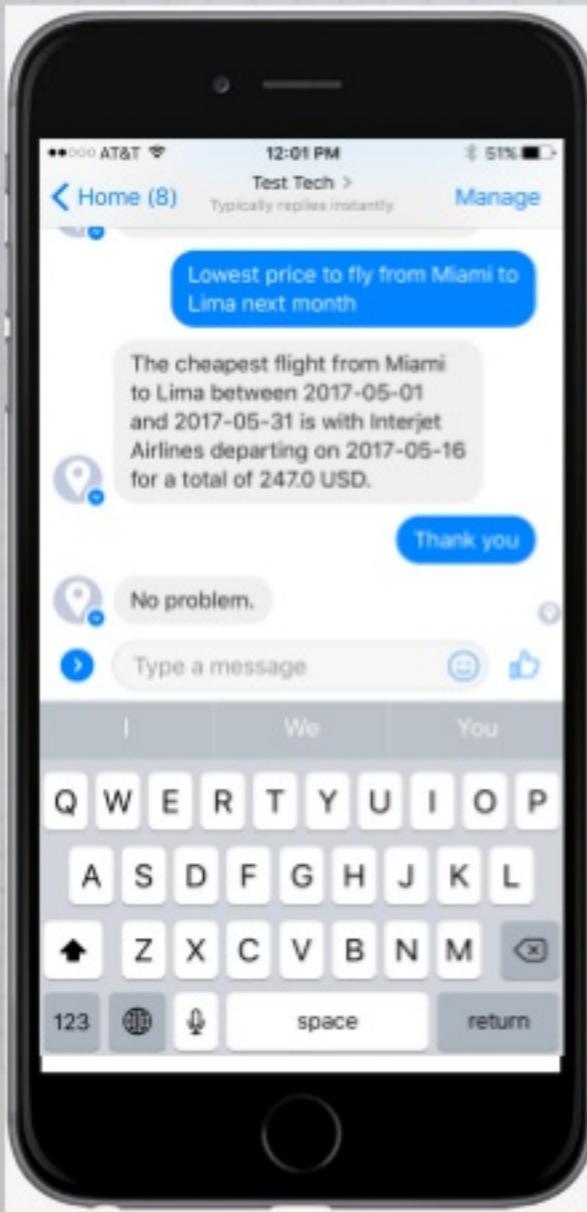
香港机票价格

仅显示部分可用及预订的单程机票价格。 请注意，一些票价适用于多个航班的预订。

上海 (SHA) 飞往 香港 (HKG) 出发日期 20/06/2017, 返回日期 25/06/2017	从 CNY770
杭州 (HGH) 飞往 香港 (HKG) 出发日期 18/06/2017, 返回日期 21/06/2017	从 CNY1,020
南京 (NKG) 飞往 香港 (HKG) 出发日期 18/06/2017, 返回日期 21/06/2017	从 CNY950
上海 (PVG) 飞往 香港 (HKG) 出发日期 11/06/2017, 返回日期 14/06/2017	从 CNY770
北京 (PEK) 飞往 香港 (HKG) 出发日期 22/06/2017, 返回日期 25/06/2017	从 CNY830
成都 (CTU) 飞往 香港 (HKG) 出发日期 22/06/2017, 返回日期 27/06/2017	从 CNY970

Datum / Date: 6-7-17

EXPLORING CHATBOTS TO BOOK FLIGHTS

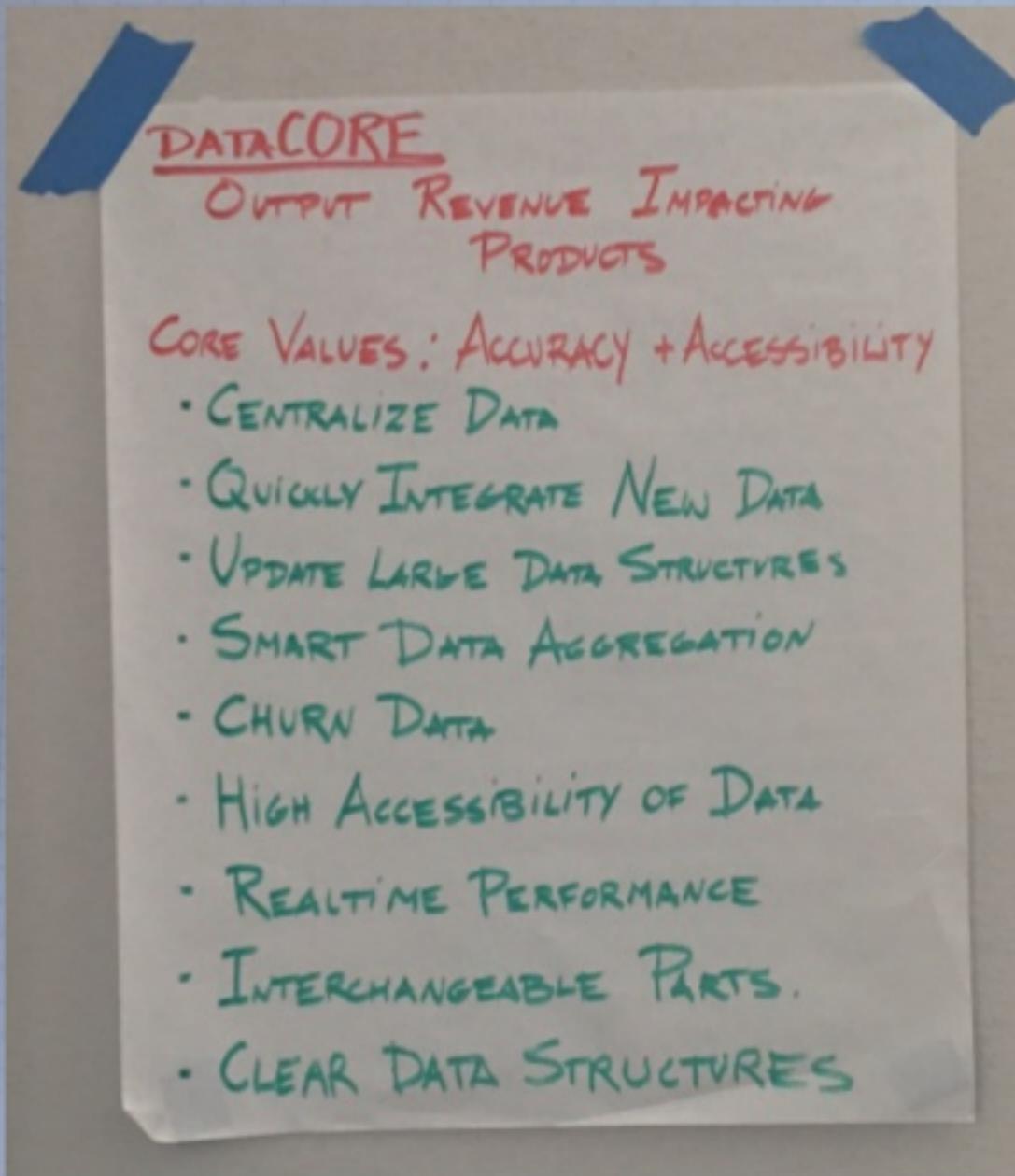


Datum / Date: 6.7.17

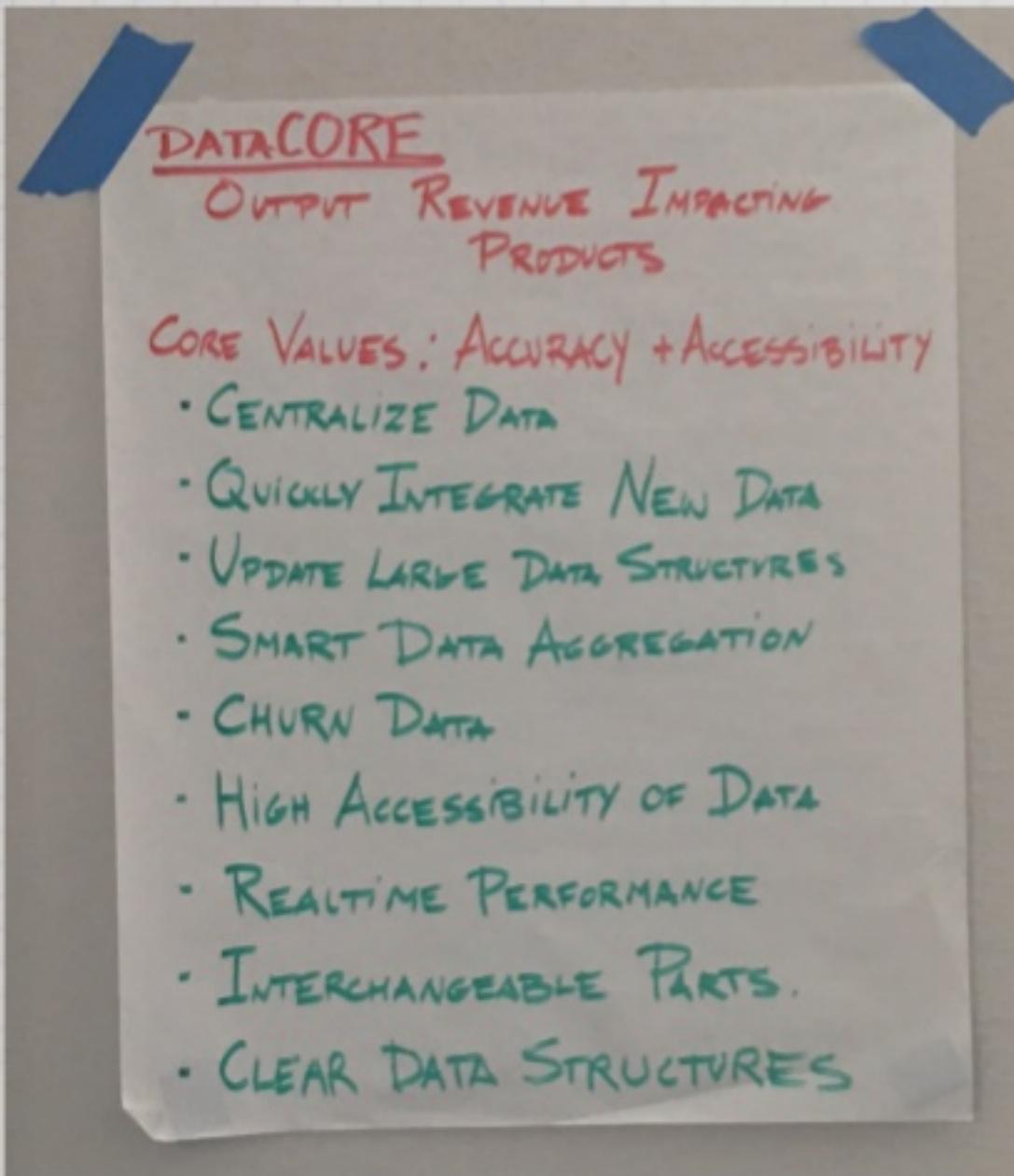
OUR CHALLENGES . . . A FEW

- VERY LARGE DATA VOLUMES
- CONCURRENT REQUESTS
- VOLATILE TICKET PRICES BASED ON AVAILABILITY
- US DOT & INTERNATIONAL ADVERTISING PRICING REGULATIONS
- MARKET & LANGUAGE DIVERSITY
- VERY DIVERSE DATA SOURCES

IN 2016 WE SET OUT TO BUILD **DATACORE**

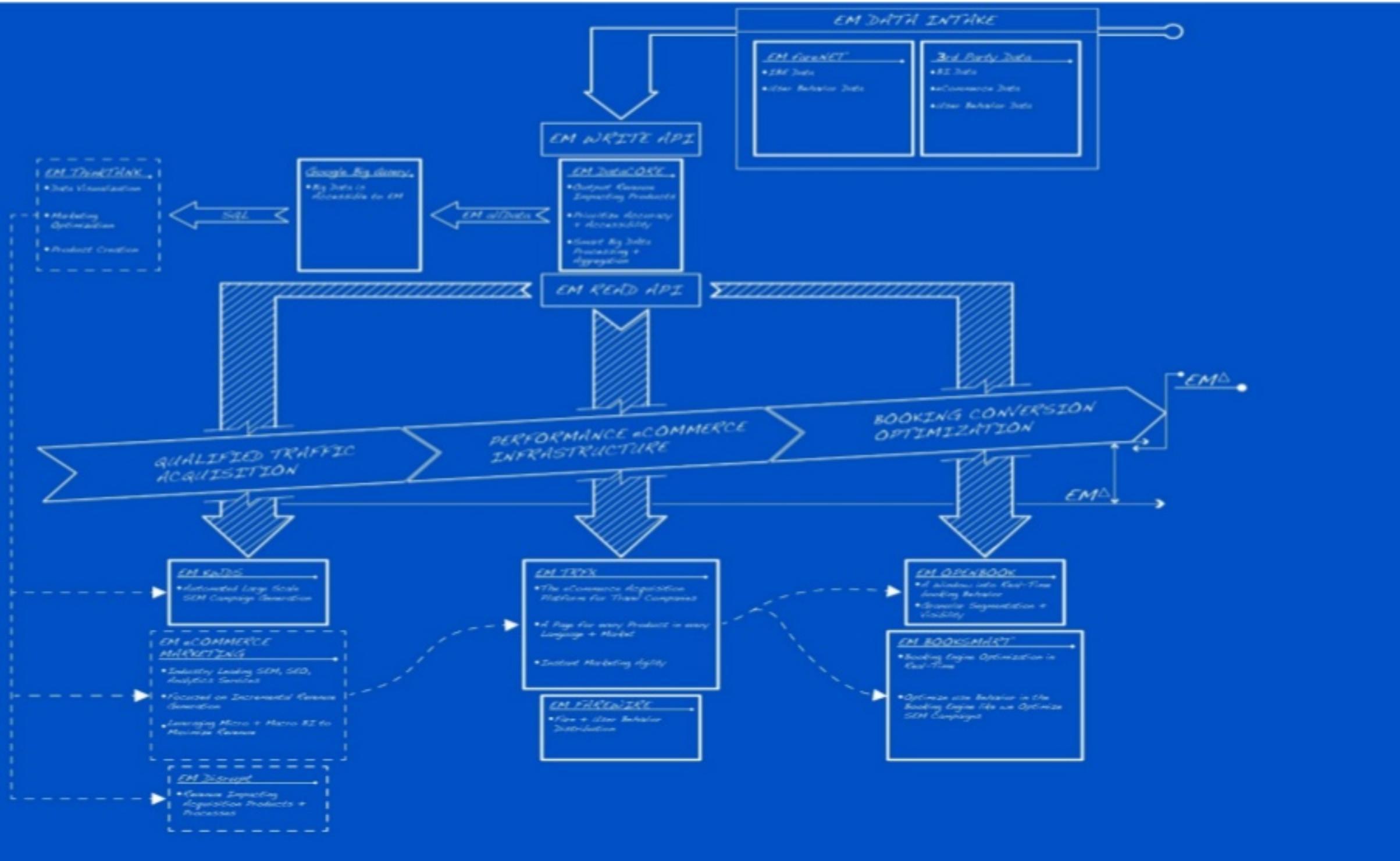


IN 2016 WE SET OUT TO BUILD **DATACORE**

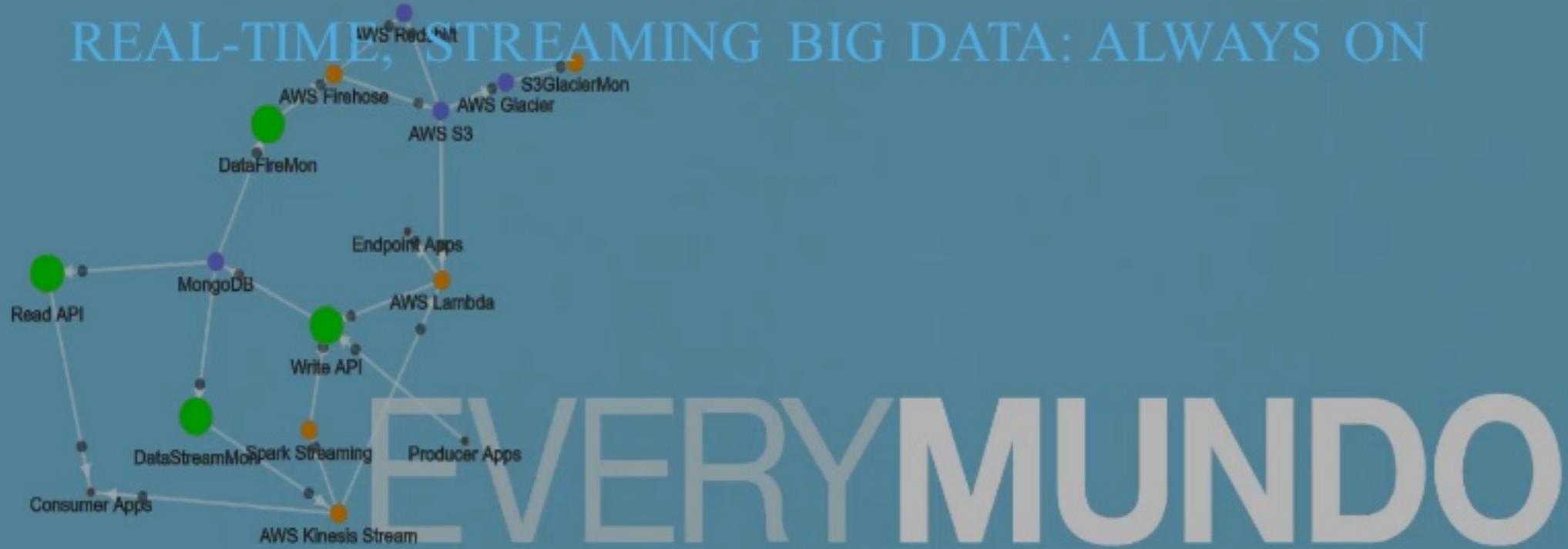


TODAY **DATACORE + Spark**

- BUILT LEVERAGING SPARK STREAMING & BATCH PROCESSING
- POWERS ALL OUR PRODUCTS
- ALLOWS OUR COMPANY TO ACCESS ALL PROCESSED DATA & CREATE PRODUCTS



REAL-TIME, STREAMING BIG DATA: ALWAYS ON



EVERYMUNDO



© 2017 EveryMundo

EVERYMUNDO

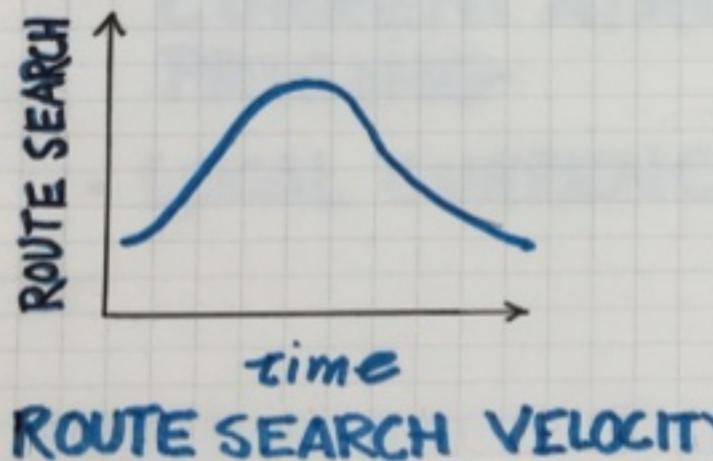
2 EXAMPLES OF Spark IN DATACORE

1-AIRINDEX & Spark

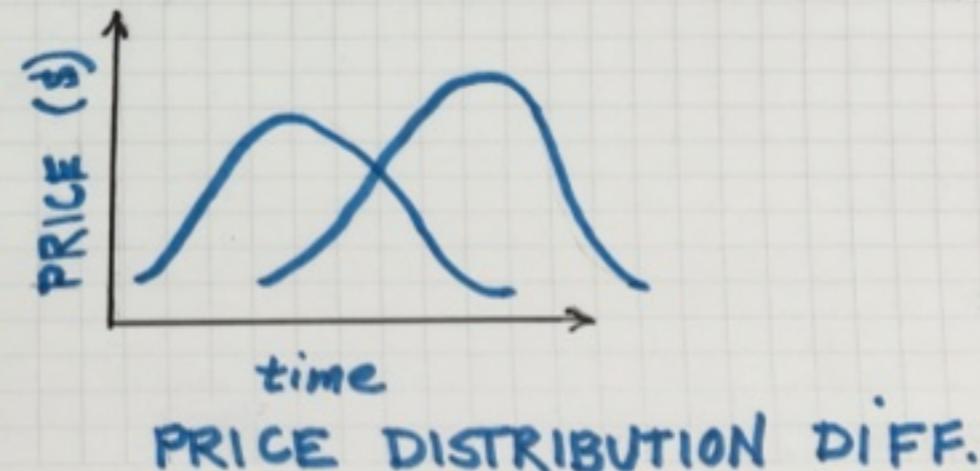
- METRIC FOR EACH AIRLINE'S ROUTE
- ANSWERS -

HOW AGGRESSIVELY SHOULD AN AIRLINE
MARKET A GIVEN ROUTE NOW AND
IN THE FUTURE

- BASED ON ACTUAL & PREDICTED



ϵ



2. AIR DPI & Spark

DYNAMIC PRICE INSERTION IN ADS

- LOGIC

- FARES COLLECTED IN REAL TIME
- BEST FARE IN 90 DAYS
- PRICES MUST MATCH CURRENT AIRLINE PRICING
- LOCAL CURRENCY

from miami to cali

All Flights Maps News Shopping More Settings Tools

About 3,440,000 results (0.84 seconds)

Fly Miami to Cali - \$404 with Copa Airlines® - copaair.com
Ad copaair.com/bookings/MIA/CLO ▾
Travel in the Americas with fast connections & full service experience onboard.
Latin America On-time Performance Service Winner 2015 – FlightStats

Miami to Cali Cheap Flights - Our Lowest Price Guarantee
Ad www.cheapoair.com/MIA-CLO/Cheap-Flights ▾
4.7 ★★★★★ rating for cheapoair.com
Grab Our Cheapest Airfares to Cali. Book Your Flights Today & Save Huge!
24/7 Customer Care · Fastest Way to Compare · Easy Booking · We Make it Easy to Travel
Ratings: Selection 9/10 - Fees 8.5/10 - Prices 8/10

Cheap One Way Flights Top 25 Flight Deals
International Flight Deal Round Trip Flights



Spark is integral to airDPI & airIndex

AIRDPI

Real-time Dynamic Price Insertion Into Ads

Changes pricing values on all online ads as ticket prices and currency exchanges fluctuate



© 2017 EveryMundo

EVERYMUNDO

The challenge

Input thousands of records/s and output
hundreds of records/s

Run the analysis 24 x 7

Ads have to be updated within a few
seconds



© 2017 EveryMundo

EVERYMUNDO

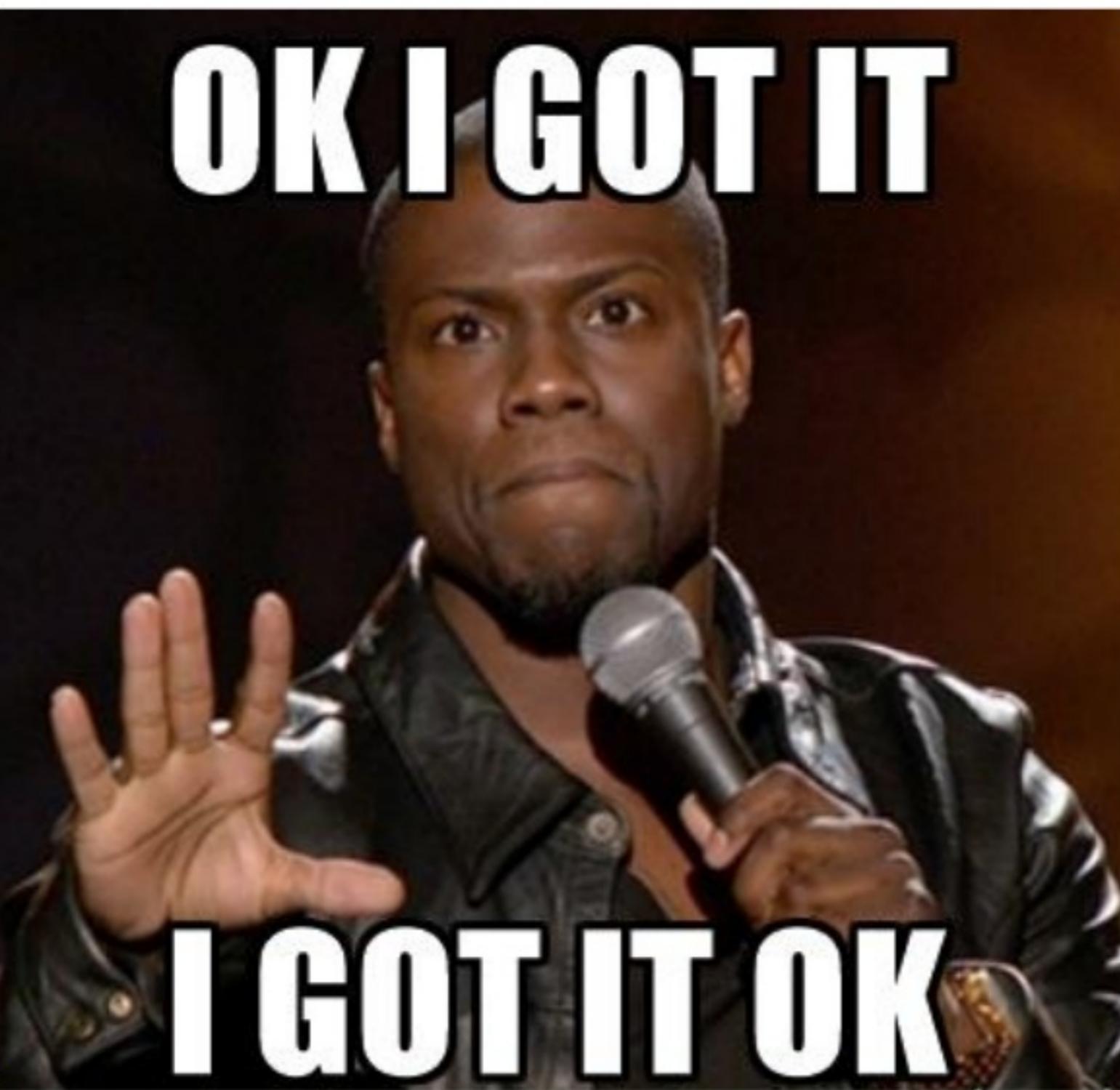
**WHAT
TO  DO ?**



© 2017 EveryMundo

EVERYMUNDO

OK I GOT IT



I GOT IT OK



© 2017 EveryMundo

EVERYMUNDO

We need a fast streaming app



© 2017 EveryMundo

EVERYMUNDO

Options

1. Starting from scratch
2. Research existing options in the market



© 2017 EveryMundo

EVERYMUNDO



STARTING FROM

SCRATCH

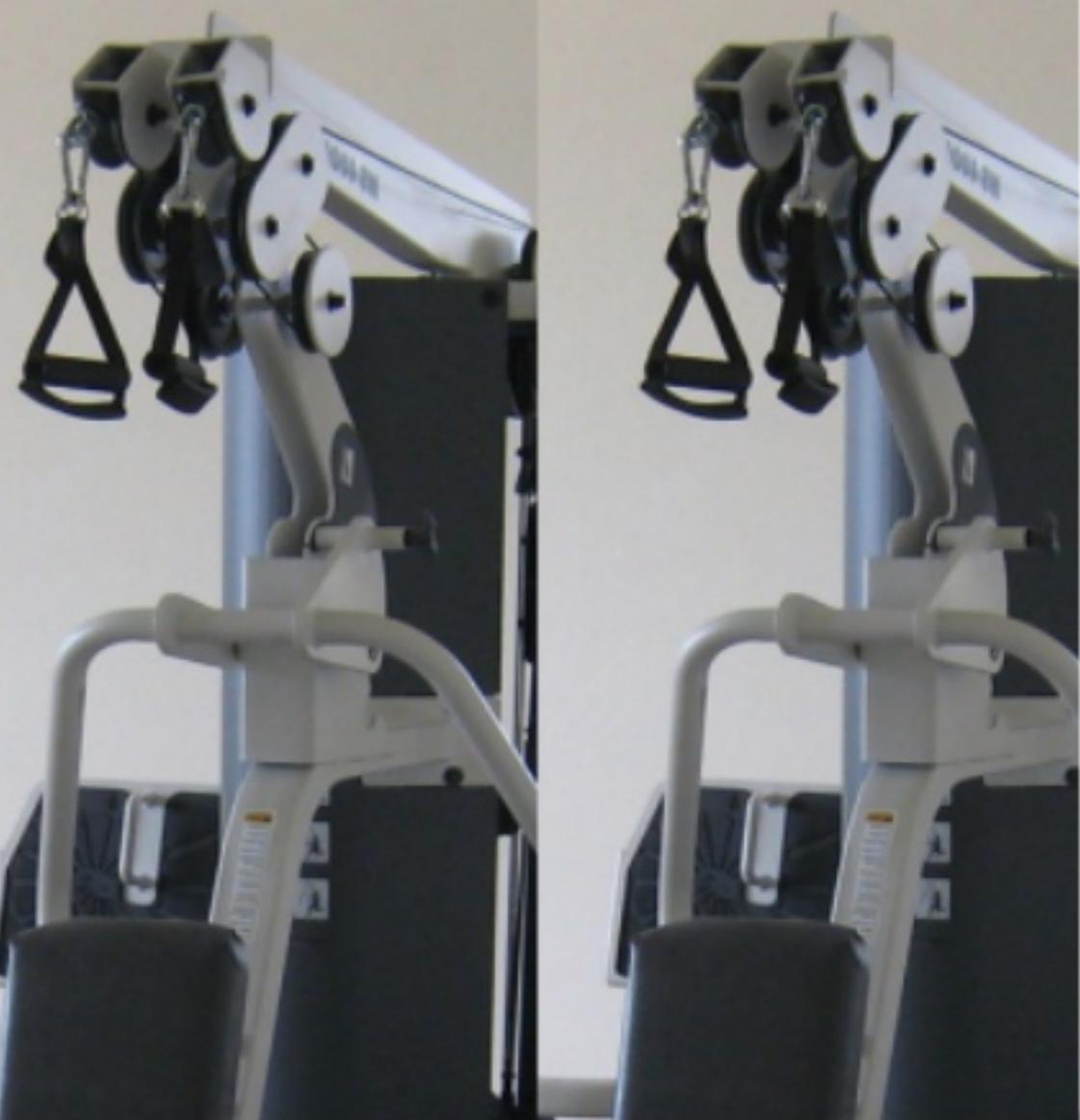


© 2017 EveryMundo

EVERYMUNDO



**WILL IT BE
EASY?
NOPE.
WORTH IT?
ABSOLUTELY.**





WILL IT BE
EASY?
NOPE.
WORTH IT?
ABSOLUTELY.

NOPE

Options



Google Cloud Dataflow



Apache Storm



Pros

- Support for many programming languages

Cons

- Complex to implement our needs

Apache Flink



Pros

- Meant for streaming
- Fast

Cons

- Only Java API
- Small Community

Google Dataflow



Pros

- Multi-framework Spark & Flink

Cons

- Only Java API (at the time)
- Small community

Apache Spark



Pros

- Multi-language
- Large community

Pros

- Fast development
- EMR and Databricks





STARTING FROM
SCRATCH



© 2017 EveryMundo

EVERYMUNDO



© 2017 EveryMundo

EVERYMUNDO

AIRDPI (Dynamic Price Insertion)

Spark 1.5

Python

updateStateByKey



© 2017 EveryMundo

EVERYMUNDO

AIRDPI (Dynamic Price Insertion)

~~Spark 1.5~~ Spark 1.6

Python

updateStateByKey



© 2017 EveryMundo

EVERYMUNDO

AIRDPI (Dynamic Price Insertion)

Spark 1.6

Python

updateStateByKey `mapWithState`



© 2017 EveryMundo

EVERYMUNDO

AIRDPI (Dynamic Price Insertion)

Spark 1.6

~~Python~~

mapWithState



© 2017 EveryMundo

EVERYMUNDO

AIRDPI (Dynamic Price Insertion)

Spark 1.6

~~Python~~ Scala 2.10

mapWithState



© 2017 EveryMundo

EVERYMUNDO

AIRDPI (Dynamic Price Insertion)

Spark 1.6.1

Scala 2.10

mapWithState



© 2017 EveryMundo

EVERYMUNDO

AIRDPI (Dynamic Price Insertion)

Spark 2.0

Scala 2.10

mapWithState



© 2017 EveryMundo

EVERYMUNDO



The Future

NEXT EXIT

AIRDPI (Dynamic Price Insertion)

Spark 2.2

Structured Streaming

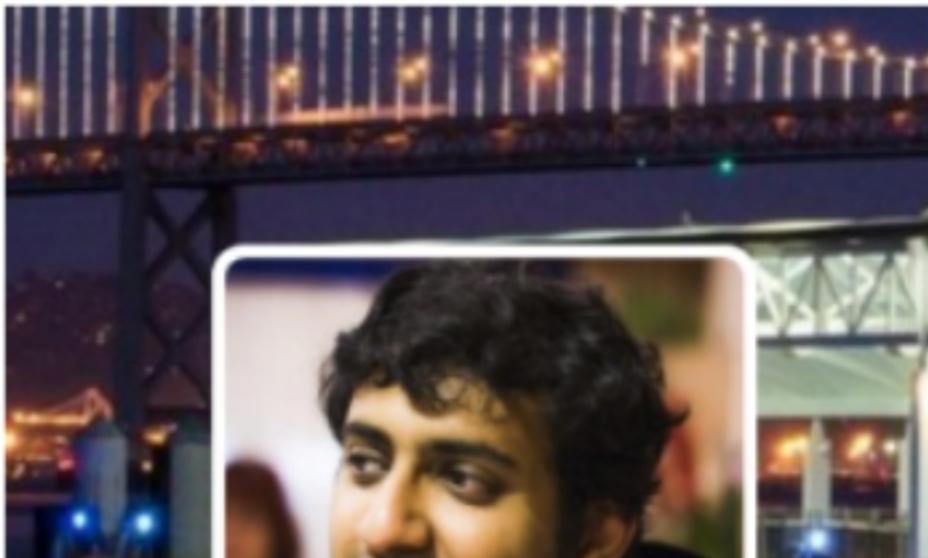
mapGroupsWithState



© 2017 EveryMundo

EVERYMUNDO

Thanks @tathadas



Tathagata Das

@tathadas

All things streaming in [@ApacheSpark](#) |
software engineer at [@Databricks](#)



© 2017 EveryMundo

EVERYMUNDO

AIRINDEX

Proprietary metric to determine route eCommerce performance

Takes into account search demand for a particular route combined with its relative price level

It answers the question: How aggressively should an airline market a given route at any point in time



© 2017 EveryMundo

EVERYMUNDO

AIRINDEX

Spark 2.1

Scala 2.11

Spark SQL / Datasets



© 2017 EveryMundo

EVERYMUNDO

AIRINDEX

Spark 2.1

Scala 2.11

Spark SQL / Datasets



© 2017 EveryMundo

EVERYMUNDO

AIRINDEX

Spark 2.1

Scala 2.11

Spark SQL / Datasets



© 2017 EveryMundo

EVERYMUNDO

AIRINDEX

Spark 2.1

Scala 2.11

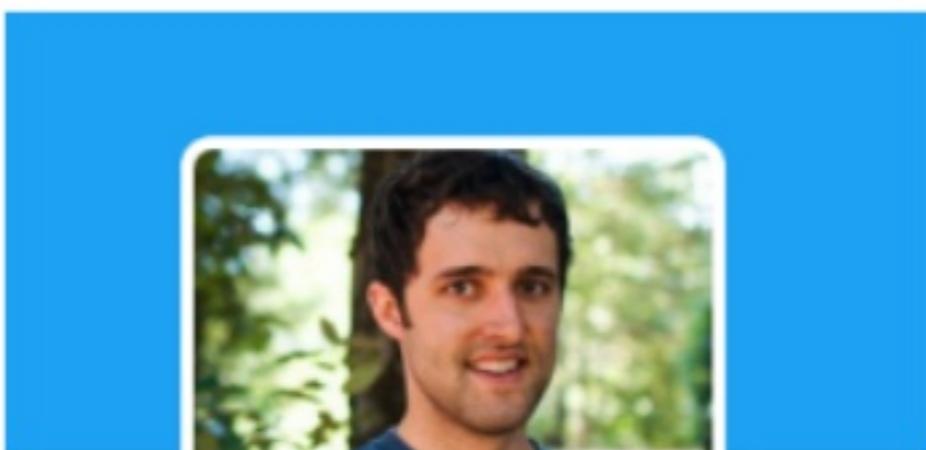
Spark SQL / Datasets



© 2017 EveryMundo

EVERYMUNDO

Thanks @michaelarmbrust



Michael Armbrust

@michaelarmbrust

Lead developer of Spark SQL

[@databricks](#), formerly [@ucberkeley](#).

Distributed databases, query languages,
scala, other nerdy stuff...

AIRINDEX

Hourly batches

Around 80,000,000 (80M) records per hour



© 2017 EveryMundo

EVERYMUNDO



~ 80 million people



© 2017 EveryMundo

EVERYMUNDO

AIRINDEX

Hourly batches

Around 80,000,000 (80M) records per hour

Around 2,000,000,000 (2B) records per day



© 2017 EveryMundo

EVERYMUNDO

~ 2 Billion

~ 1.37 Billion



~ 324 Million



~ 305 Million



~ 80 Million



AIR**INDEX**

Amazon S3 buckets

Amazon Redshift

ORC / Parquet



© 2017 EveryMundo

EVERY**MUNDO**



The Future

NEXT EXIT

AIRINDEX

Structured Streaming

Spark 2.2 mapGroupsWithState

MLlib



© 2017 EveryMundo

EVERYMUNDO

New and very useful features

Reliability

Clean API for batch streaming

Great performance

Simple Management



© 2017 EveryMundo

EVERYMUNDO

That is why we chose



© 2017 EveryMundo

EVERYMUNDO



Thank You.

www.everymundo.com

@everymundo

@danielx