Catching Signals in RM

Augment predictive Revenue Management models with contextual data

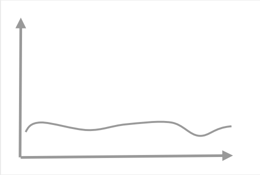
JAN 20, LHG Revenue Management & Migacore

Tabea Hasler & George Brova

Internal

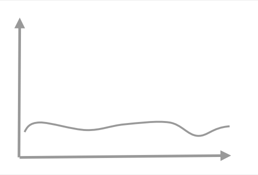
Background: Peaks & Trends influence Customer Behavior significantly

Peaks & Trends



Background: Peaks & Trends influence Customer Behavior significantly

Peaks & Trends



AMLDEPFL

AMLD at EPFL
25.01.-29.01.20

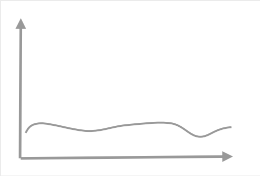
Background: Peaks & Trends influence Customer Behavior significantly

Peaks & Trends

Austrian Lufthansa SWISS LUFTHANSA GROUP

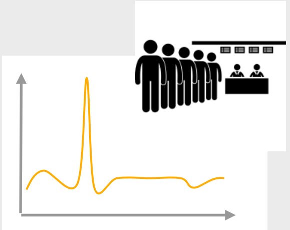
Background: Peaks & Trends influence Customer Behavior significantly

Peaks & Trends



AMLD EPFL

AMLD at EPFL
25.01.-29.01.20



Today

Human expert
(Sales/RM)

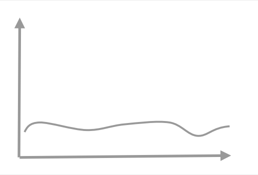
+

Historical bookings

- Manual peak and trend detection
- Manual adjustment of forecast and steering
- No use of additional sources of information

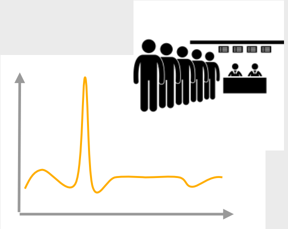
Background: Peaks & Trends influence Customer Behavior significantly

Peaks & Trends



AMLDEPFL

AMLD at EPFL
25.01.-29.01.20



Today

Human expert (Sales/RM)

+

Historical bookings

- Manual peak and trend detection
- Manual adjustment of forecast and steering
- No use of additional sources of information

PAIN

Limited information leads to **lagging behind** rather than **knowing in advance**

Idea: Include Contextual Data in our Models to detect Travel Patterns

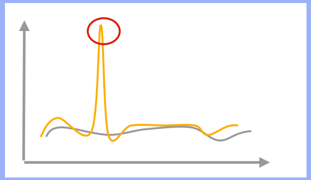
Idea: Include Contextual Data in our Models to detect Travel Patterns

New Input

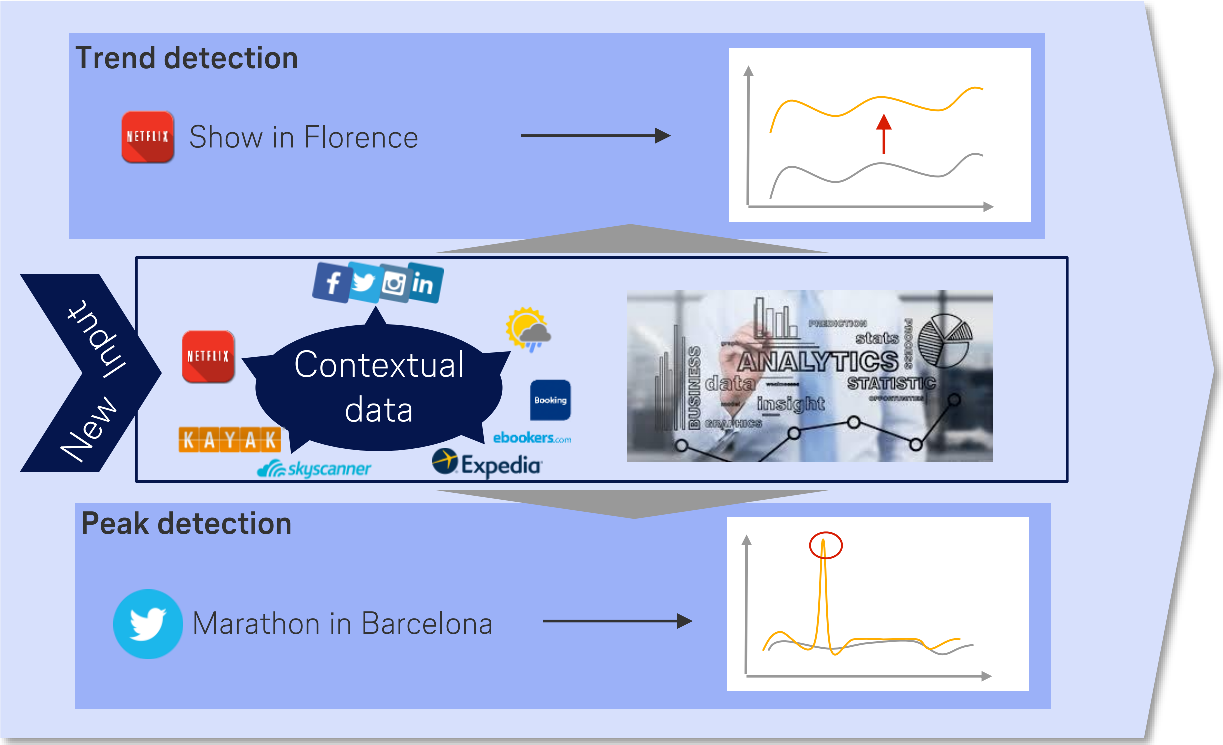
Contextual data

Peak detection

Marathon in Barcelona →

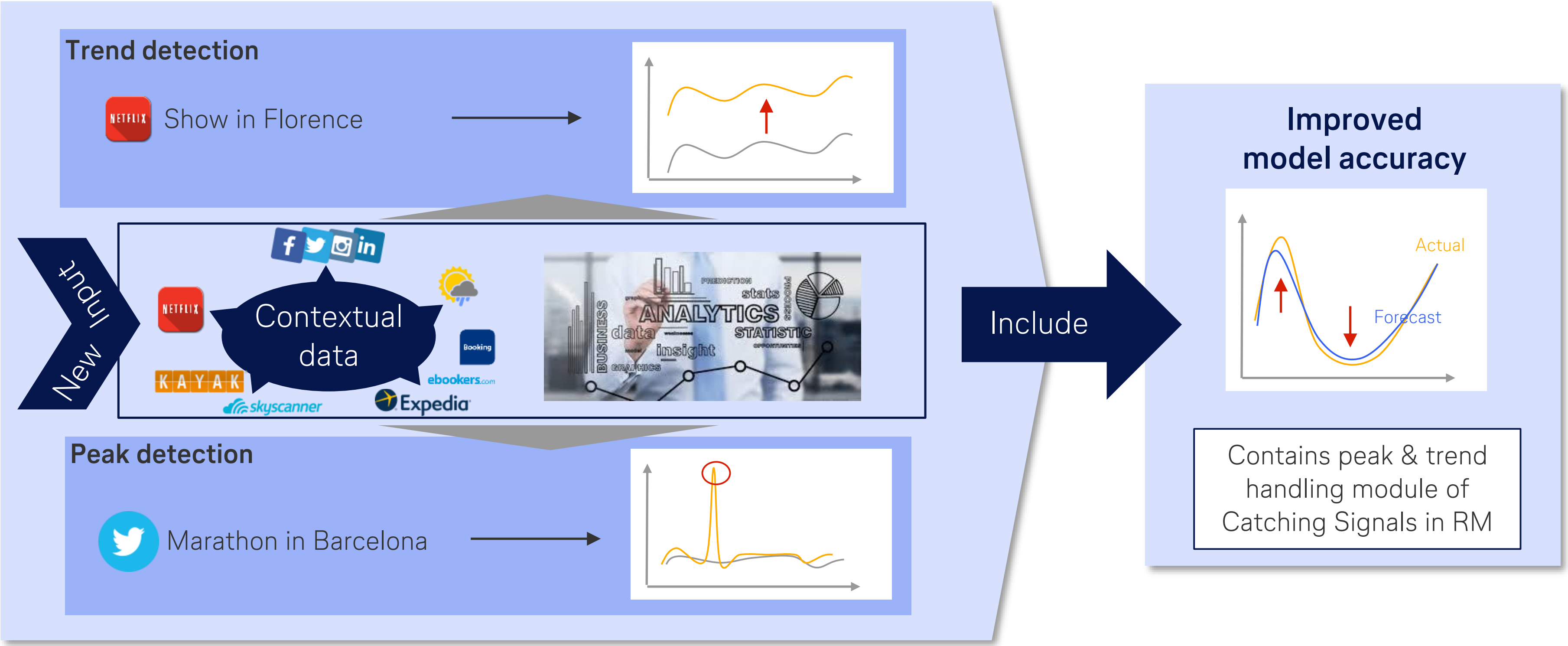


Idea: Include Contextual Data in our Models to detect Travel Patterns



Austrian Lufthansa SWISS LUFTHANSA GROUP

Idea: Include Contextual Data in our Models to detect Travel Patterns

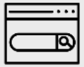


Austrian Lufthansa SWISS

LUFTHANSA GROUP

Application: Create Data Signals for direct incorporation in Models

Create Signals



Search Engines

Events / News, ...

Social media

Train Signals

Add signals to model

$$
\begin{aligned}
\log\Big(\lambda(\mathbf{x}_{i,t},t)\Big) = {} & \beta_0 + \beta_{\xi}\hat{\xi}_{i,t} + \sum_{j=1}^{6} \mathbf{1}_{\mathrm{BDAY}_{i,t}=j}\beta_{1,j} + f_p\left(p_{i,t}\right) \\
& + f_{p,t}\big(p_{i,t},t\big) + f_{p,1}\left(p_{i,t},\mathrm{DTIME}_{i,t}\right) + f_{p,2}\big(p_{i,t},\mathrm{YDAY}_{i,t}\big) \\
& + f_t\left(t\right) + f_1\left(\mathrm{DTIME}_{i,t}\right) + f_2\left(\mathrm{YDAY}_{i,t}\right) \\
& + f_{t,1}\left(t,\mathrm{DTIME}_{i,t}\right) + f_{t,2}\big(t,\mathrm{YDAY}_{i,t}\big) + f_{1,2}\big(\mathrm{DTIME}_{i,t},\mathrm{YDAY}_{i,t}\big) \\
& + f_{p,a}\big(p_{i,t},\mathrm{INDEX}_{1,i,t}\big) + f_{p,b}\big(p_{i,t},\mathrm{INDEX}_{2,i,t}\big) + \ldots \\
& + f_{p,I}\big(p_{i,t},\mathrm{SIGNAL}_{1,i,t}\big) + f_{p,II}\big(p_{i,t},\mathrm{SIGNAL}_{2,i,t}\big) + \ldots \\
& + \ldots
\end{aligned}
$$

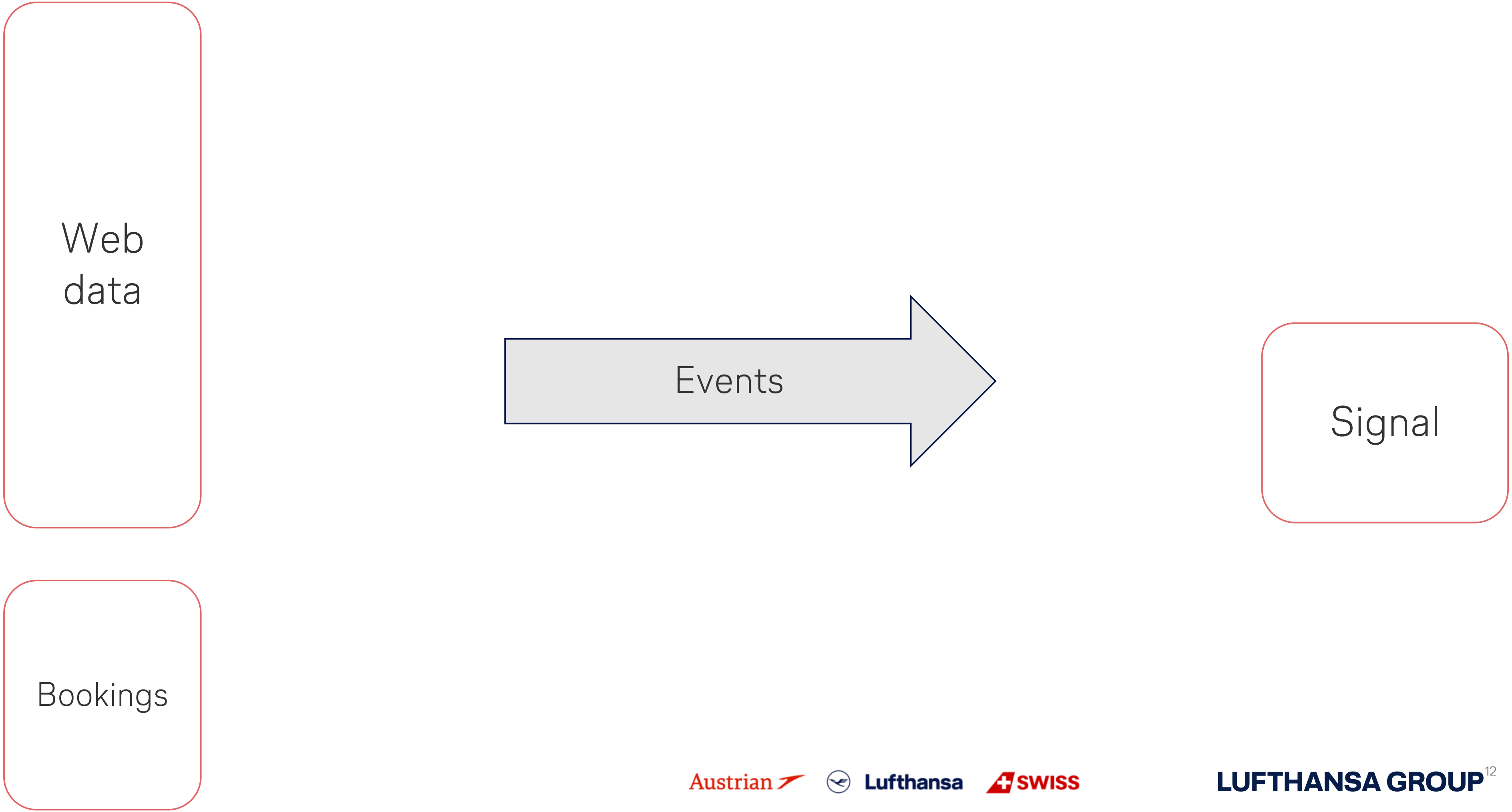
\+ additional Indexes/Signals

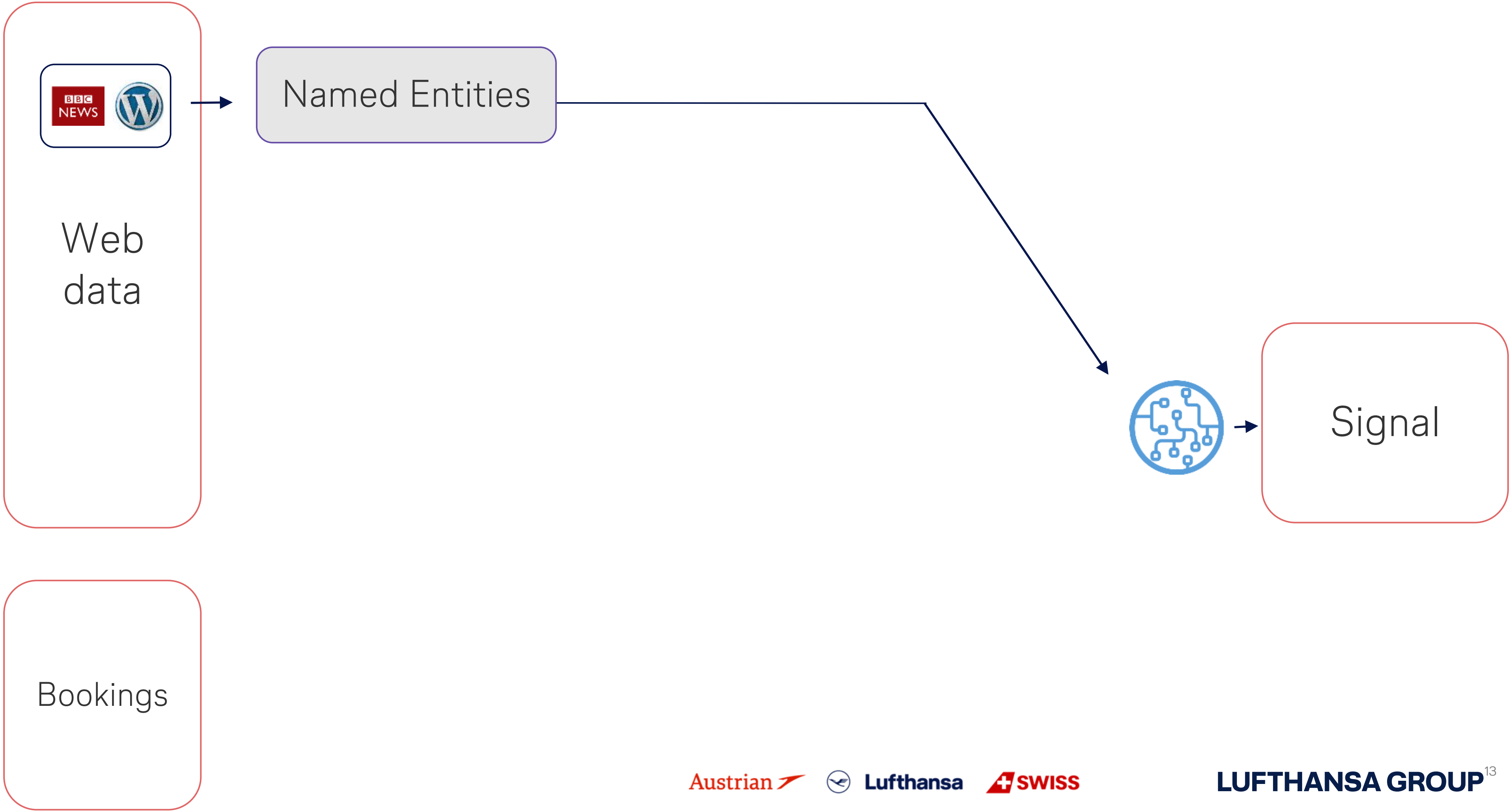
Regression Model

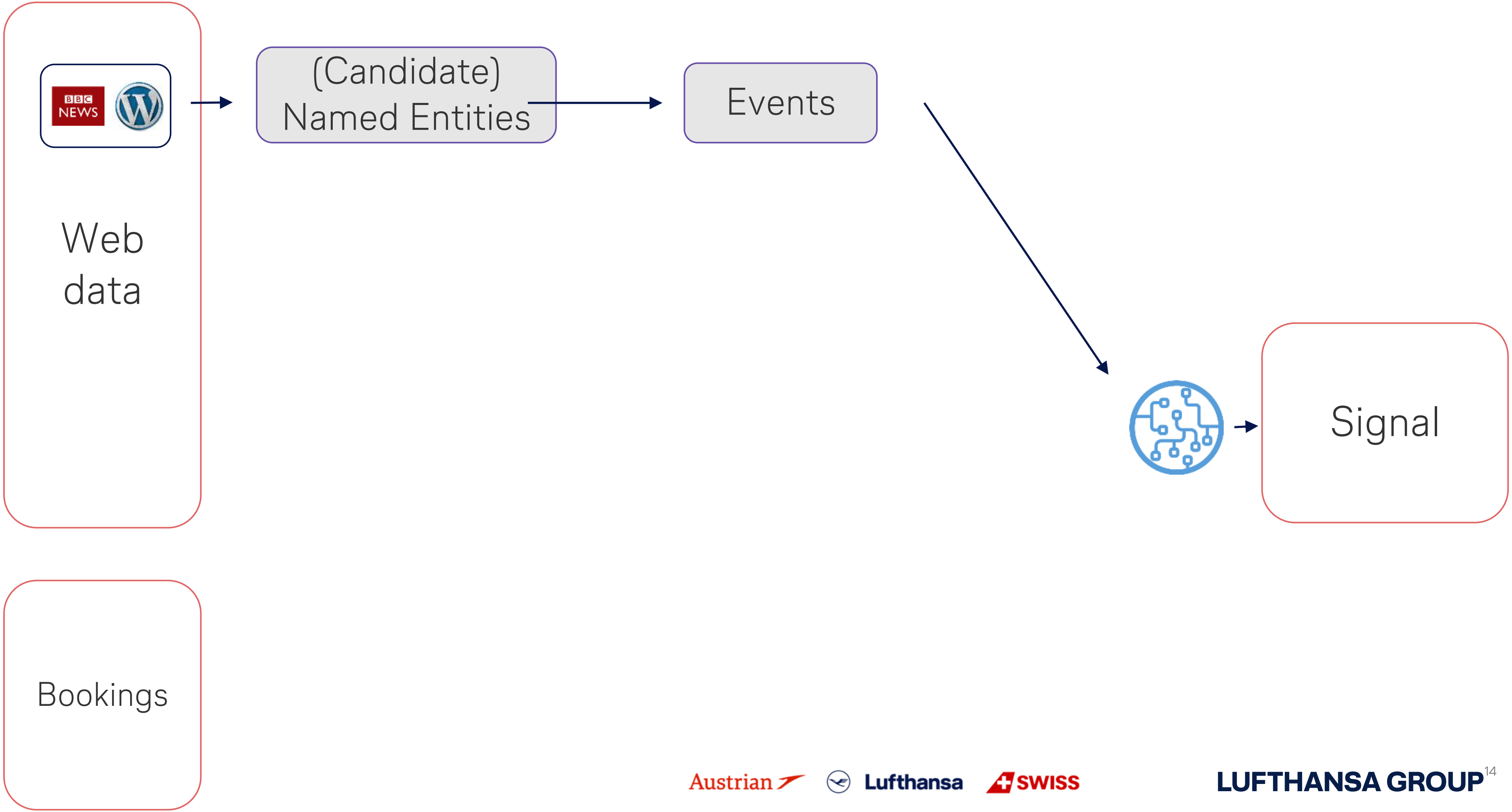
GLM/GAM

Count Data, Bookings

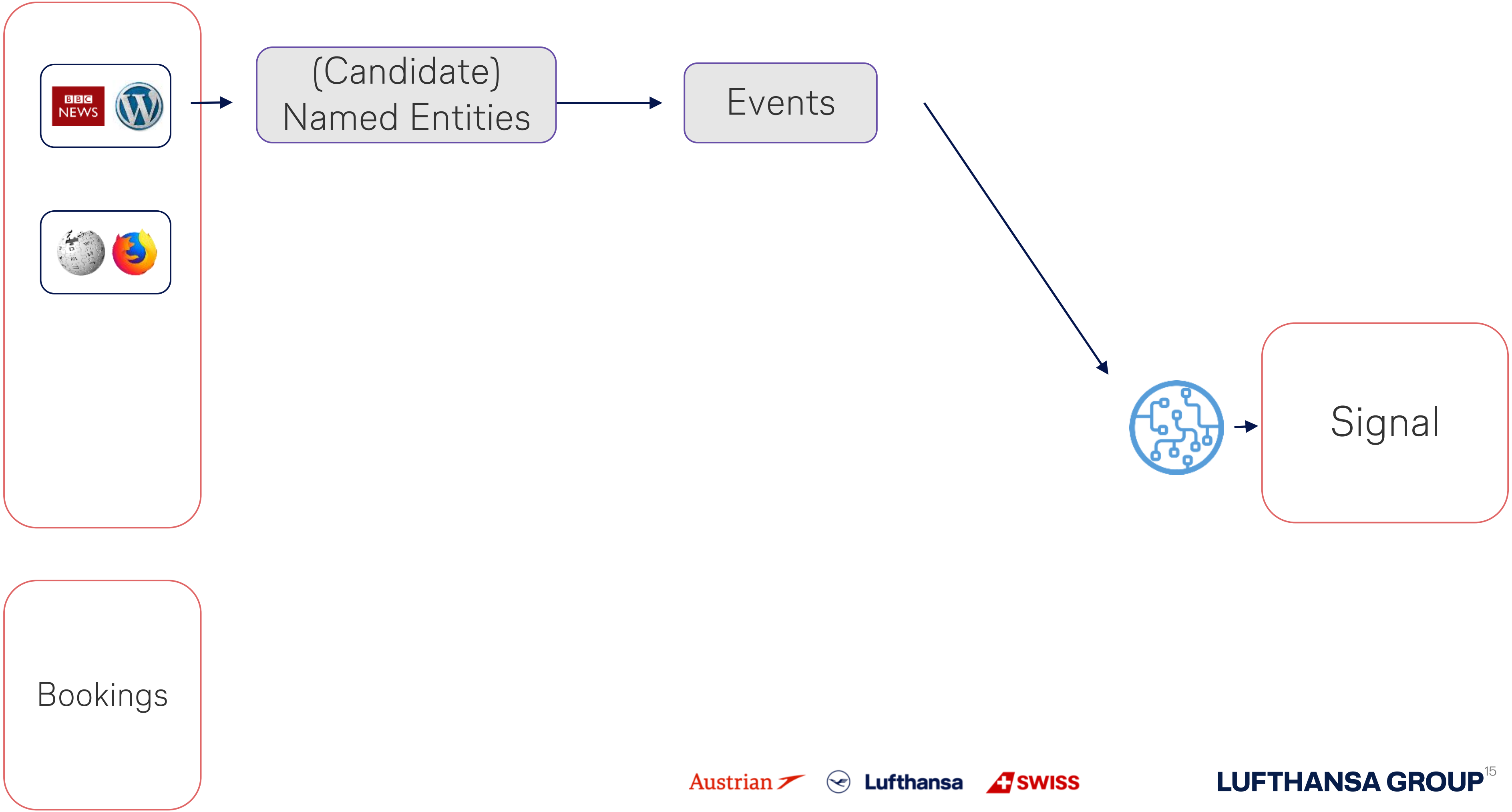
Austrian Lufthansa SWISS LUFTHANSA GROUP


Web data
Bookings
Events
Signal

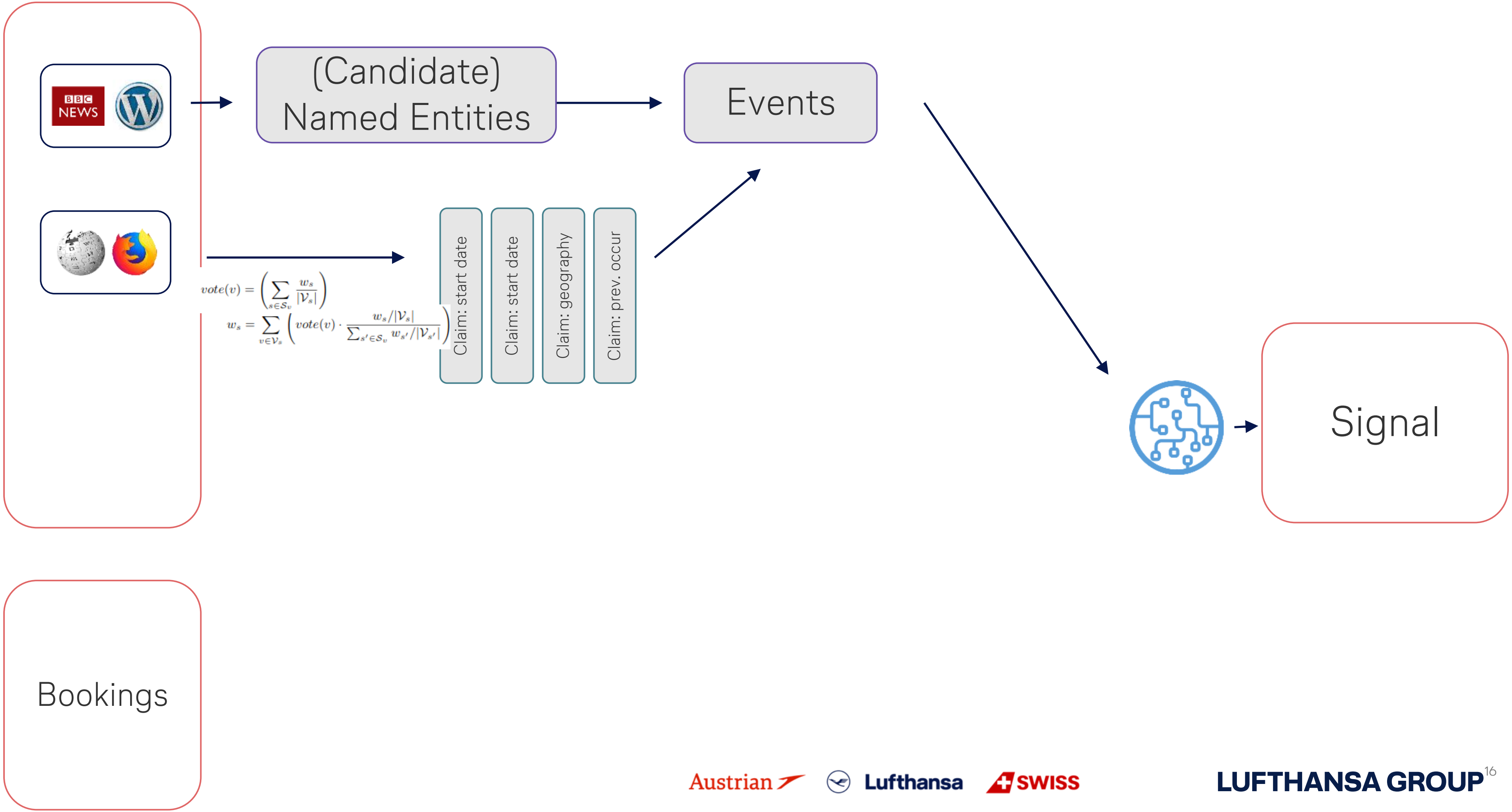

BBC NEWS
Web data
Named Entities
Signal
Bookings

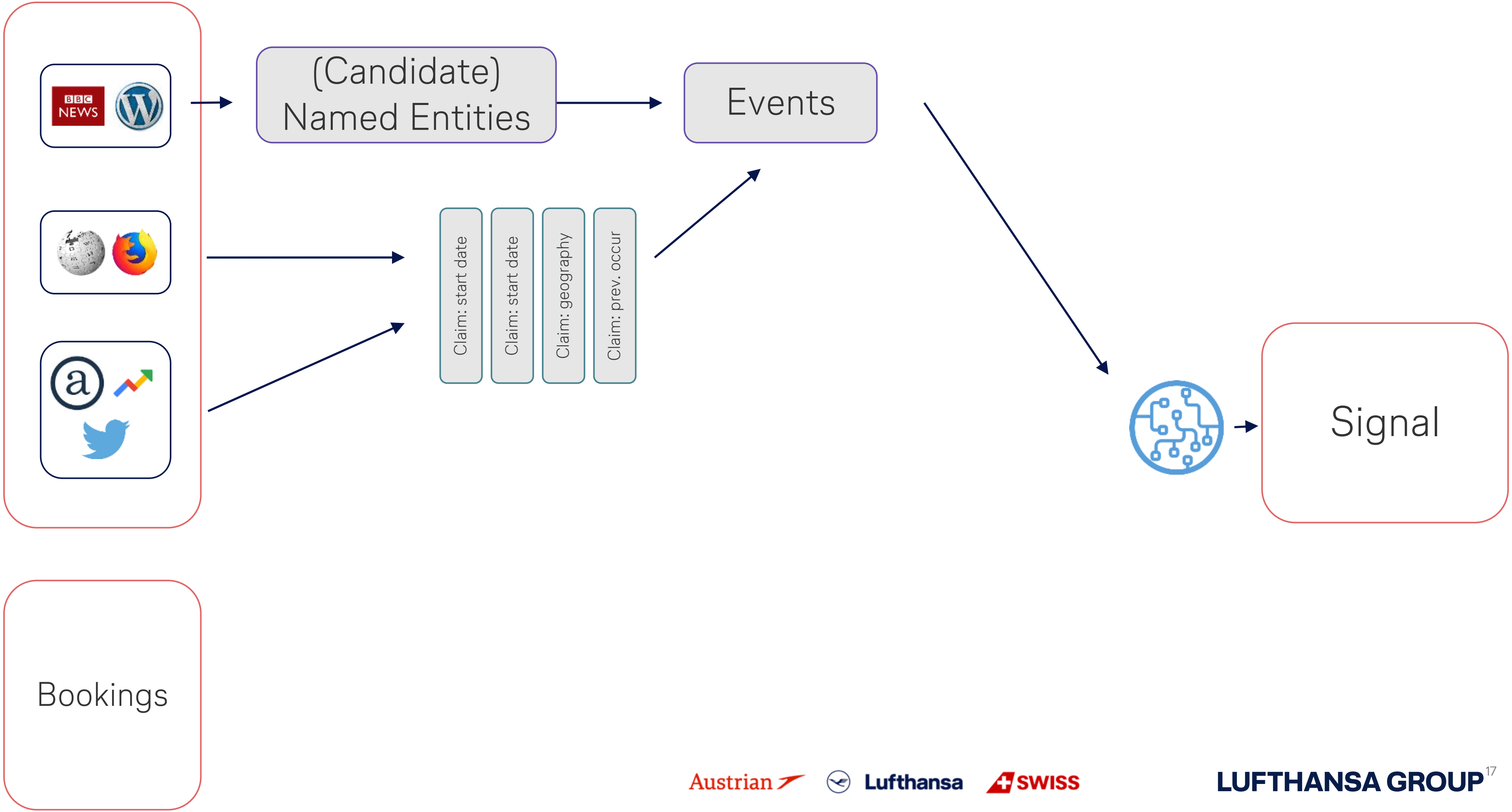


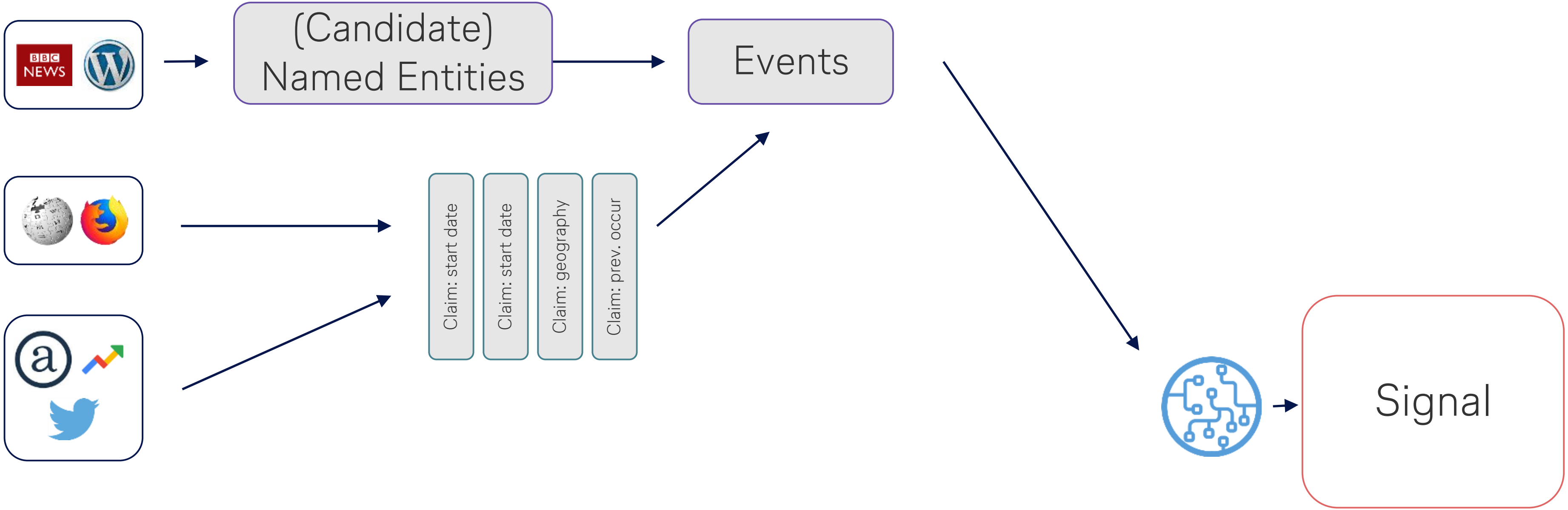
BBC NEWS
Web data
(Candidate) Named Entities
Events
Signal
Bookings

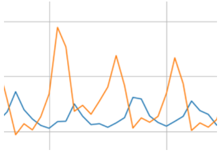


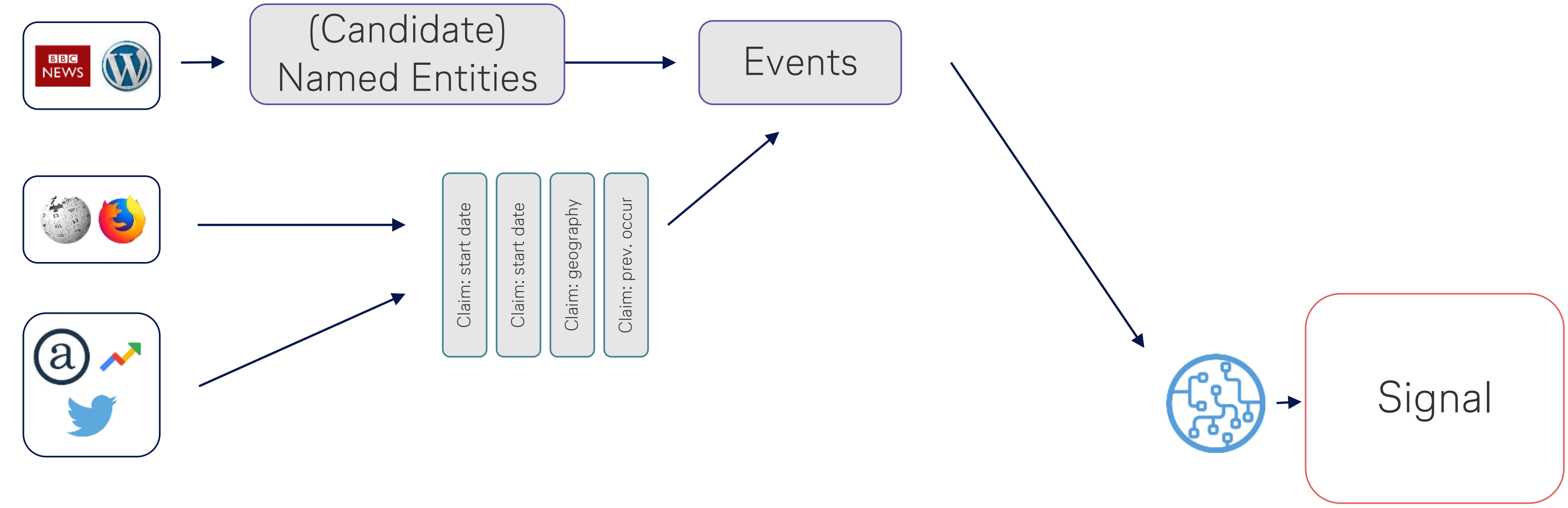
BBC NEWS
(Candidate) Named Entities
Events
Signal
Bookings


(Candidate)
Named Entities
Events
$vote(v) = \left(\sum_{s \in \mathcal{S}_v} \frac{w_s}{|\mathcal{V}_s|}\right)$
$w_s = \sum_{v \in \mathcal{V}_s} \left(vote(v) \cdot \frac{w_s/|\mathcal{V}_s|}{\sum_{s' \in \mathcal{S}_v} w_{s'}/|\mathcal{V}_{s'}|} \right)$
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
Signal
Bookings

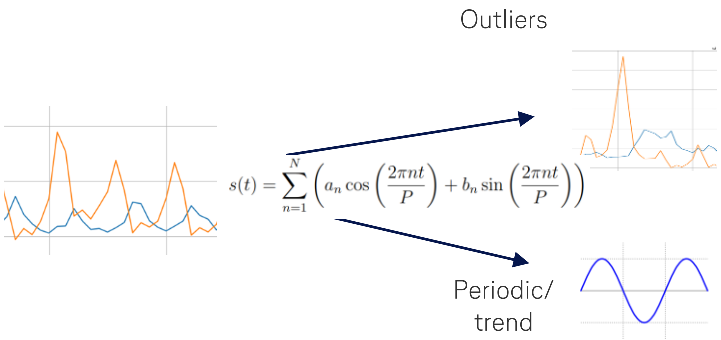

(Candidate)
Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
Signal
Bookings


BBC NEWS
(Candidate)
Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
Signal

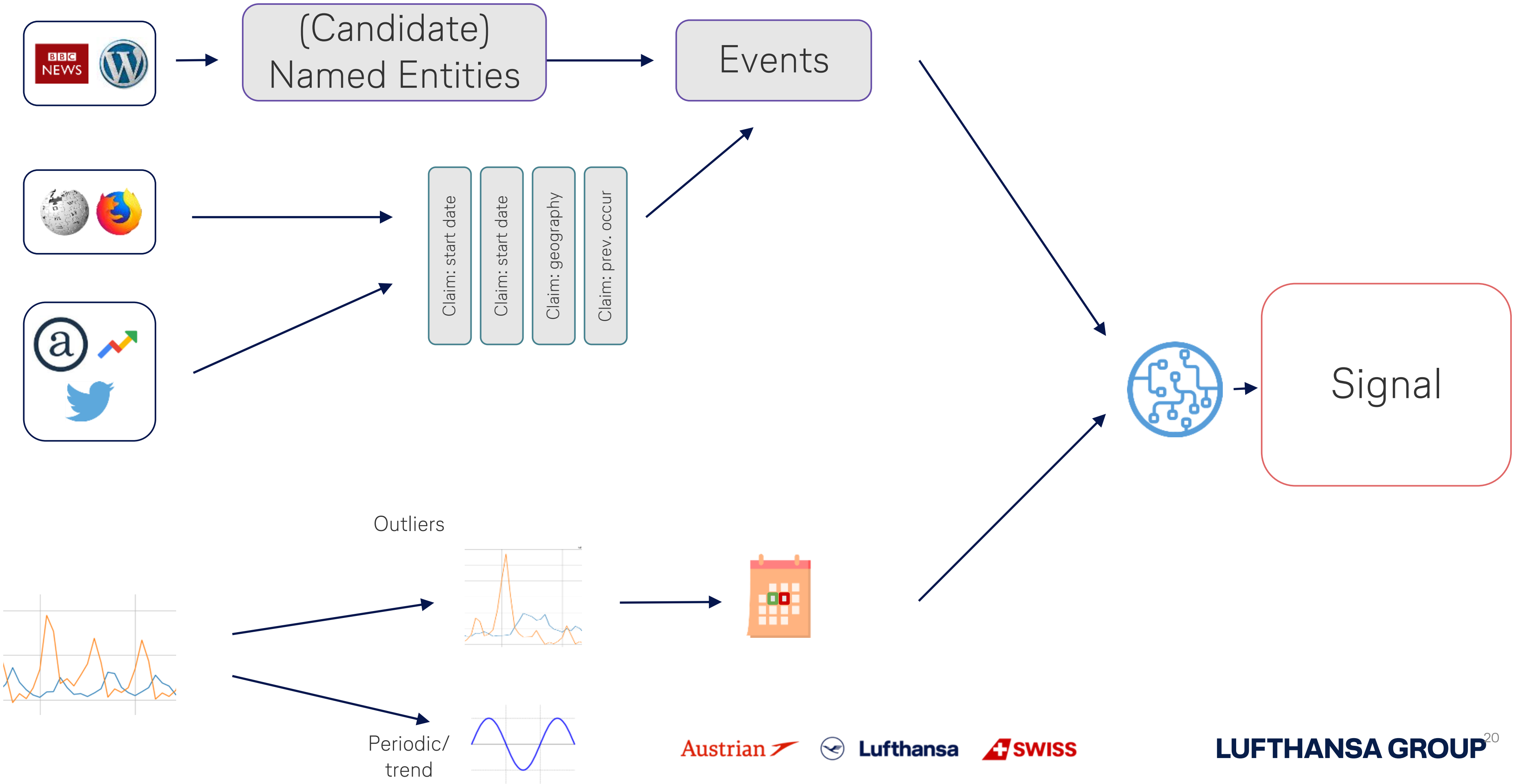


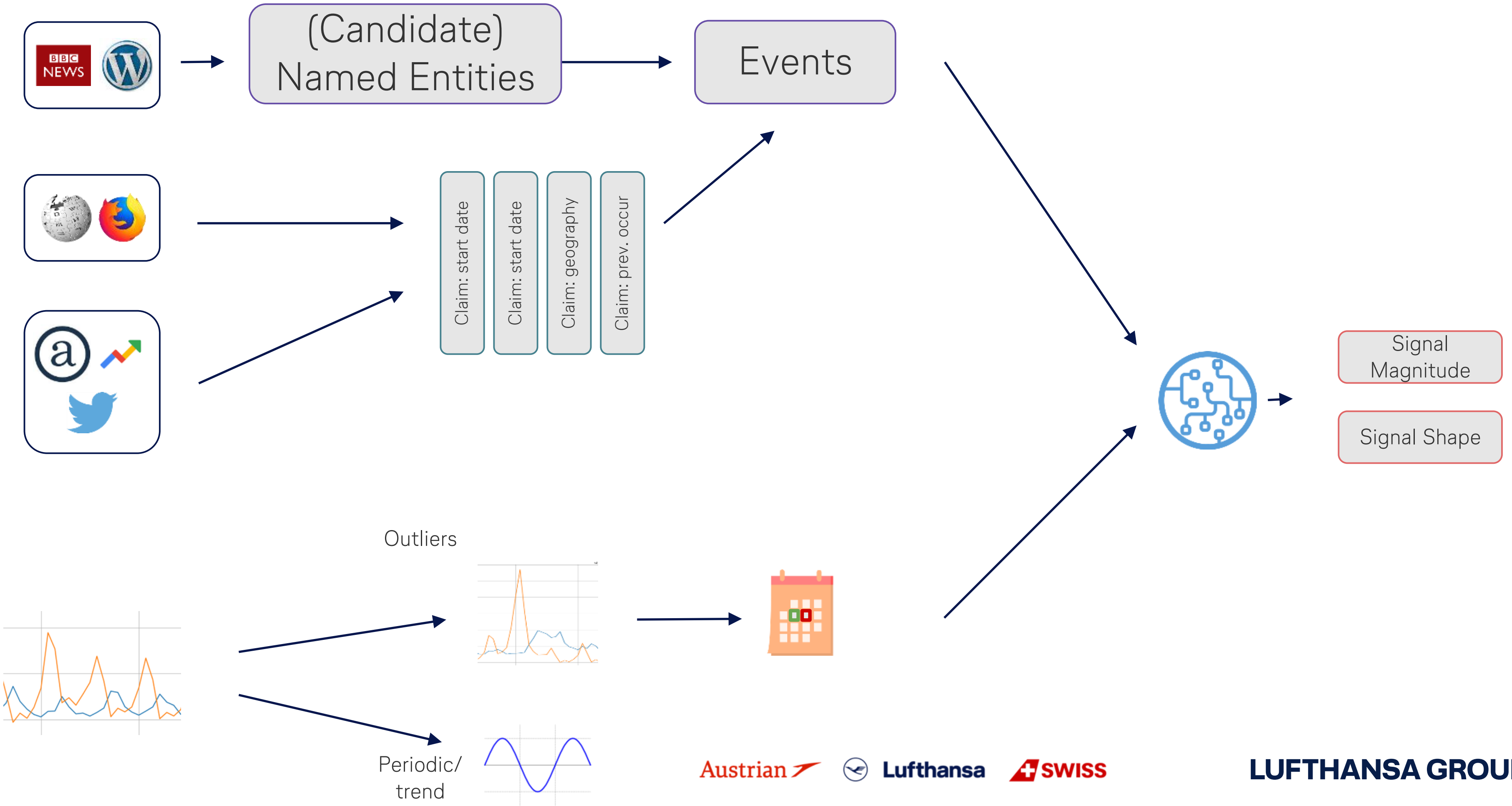


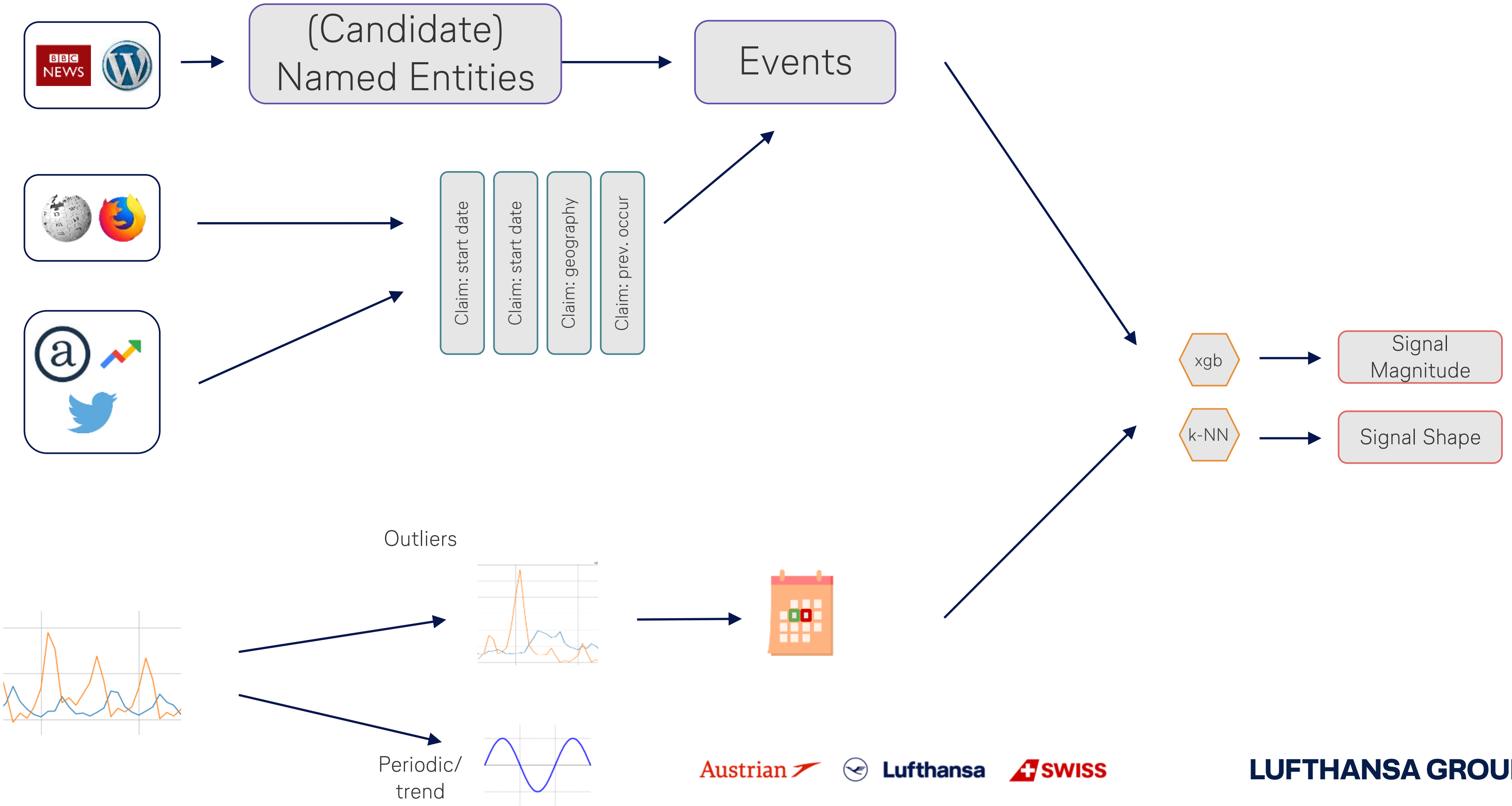
(Candidate)
Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
Signal

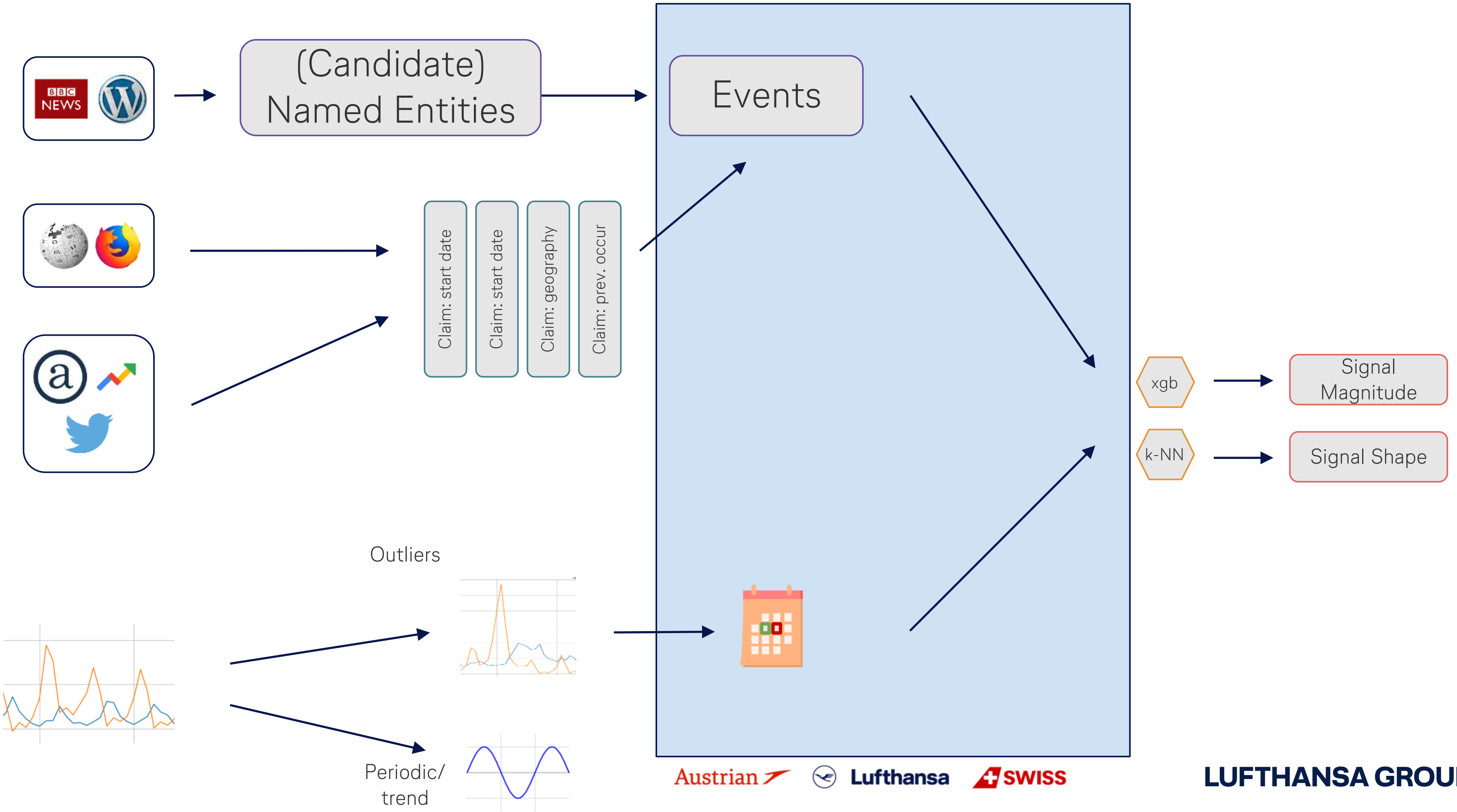


Outliers
$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right) \right)$
Periodic/
trend

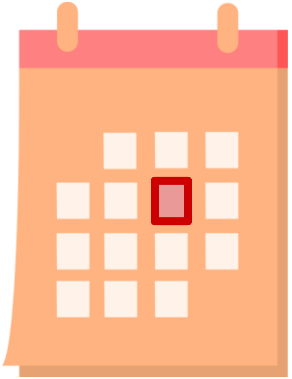

BBC NEWS
(Candidate) Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
Signal
Outliers
Periodic/ trend

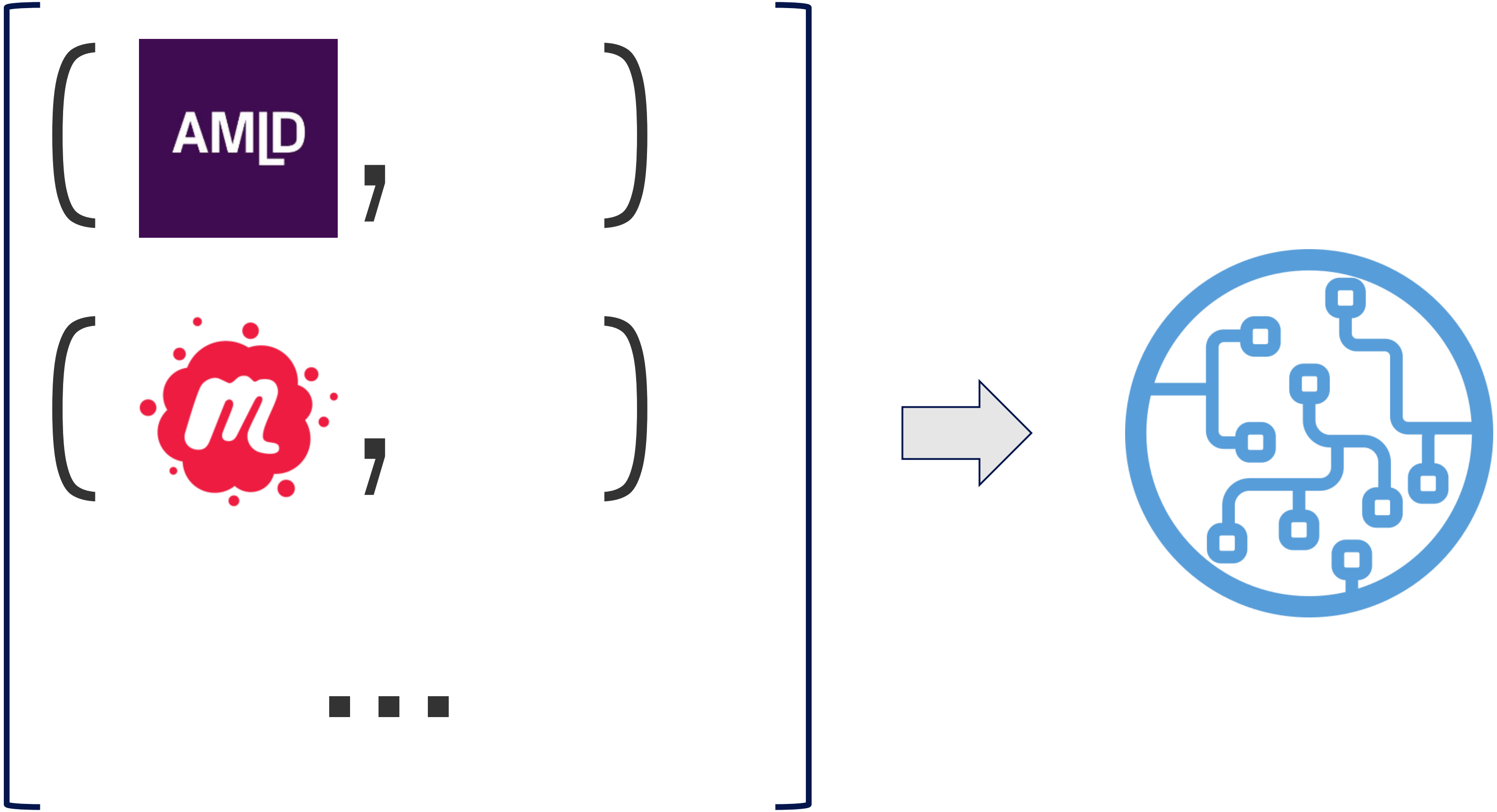

(Candidate)
Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
Signal
Magnitude
Signal Shape
Outliers
Periodic/
trend

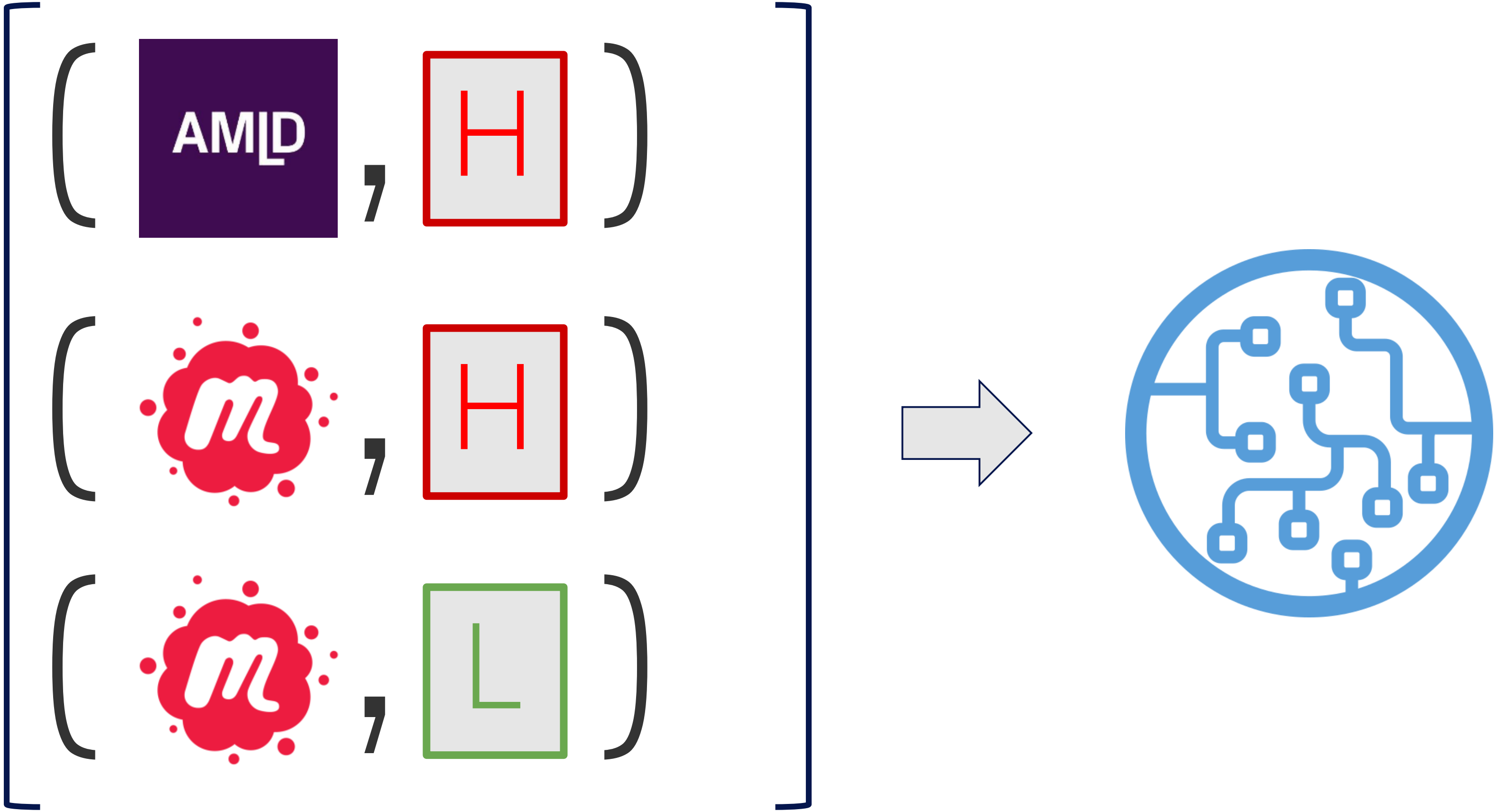

(Candidate)
Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
xgb
k-NN
Signal Magnitude
Signal Shape
Outliers
Periodic/
trend

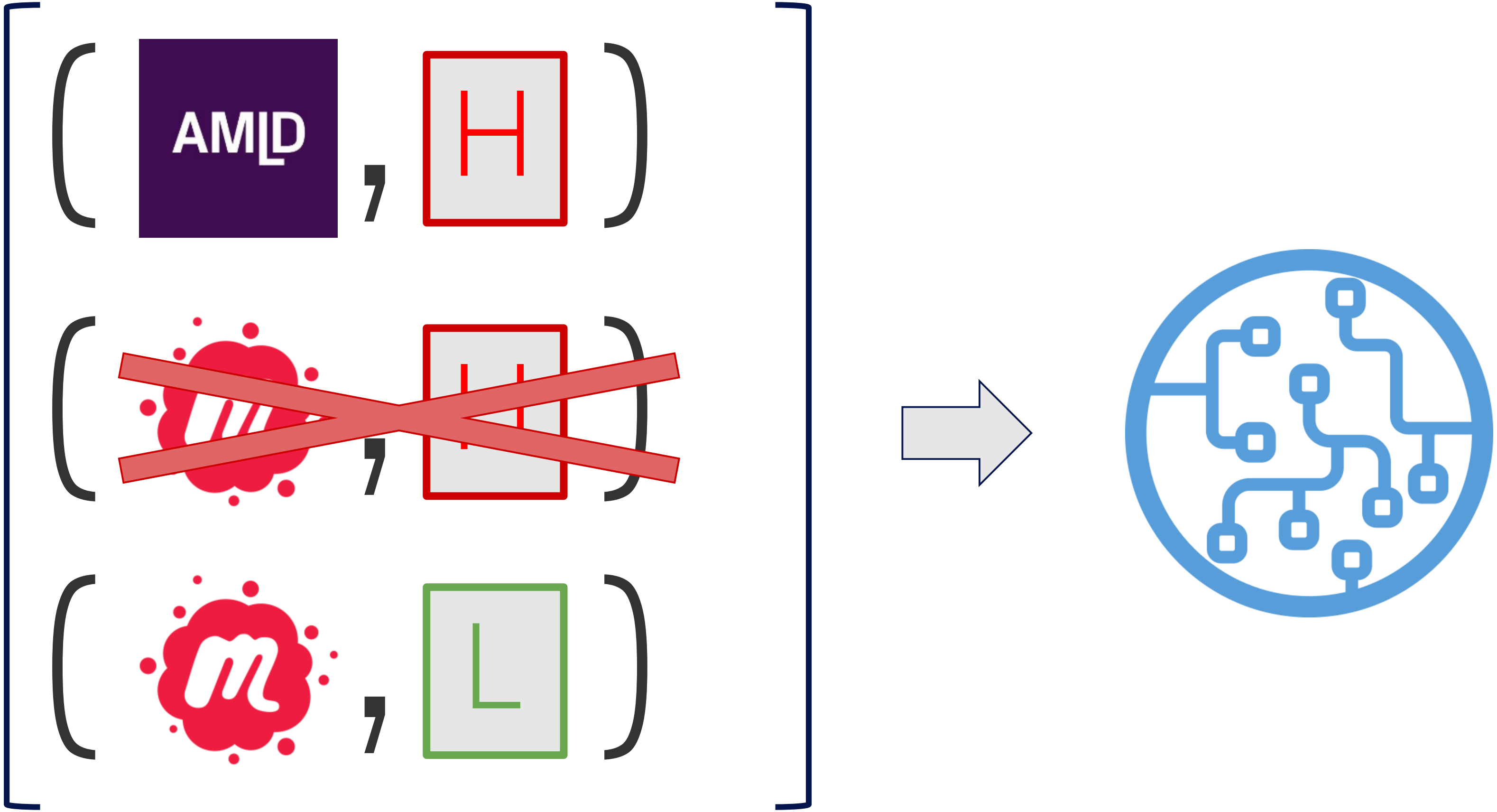

(Candidate)
Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
Outliers
Periodic/
trend
xgb
k-NN
Signal Magnitude
Signal Shape

Deep dive: how do same-day events interact?




AMLD

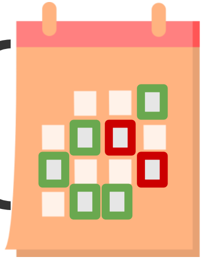

AMLD
H
H
L


AMLD
H
H
L

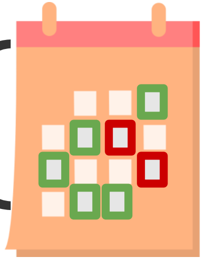
Alternative approach

Events have (independent) probability to cause impact

Set up as maximum likelihood

P() = P($\boxed{L}$) · P($\boxed{H}$) · ...

Set up as maximum likelihood

P() = P($\boxed{L}$) · P($\boxed{H}$) · ...

$$P(\text{data}) = \prod_{D \in \mathcal{D}} P(D) \prod_{\overline{D} \in \overline{\mathcal{D}}} P(\overline{D}) \qquad \text{where } P(\overline{D}) = \prod_{e_i \in \overline{D}} (1 - p_{e_i})$$

$$P(D) = 1 - \prod_{e_i \in D} (1 - p_{e_i})$$

If two events are similar, their impact probabilities should be similar

If two events are similar, their impact probabilities should be similar

$$\uparrow \mathrm{sim}(\text{AMLD}, \text{ICML}) \Rightarrow p_{\text{AMLD}} \approx p_{\text{ICML}}$$

$$\text{subject to} \sum_{i,j} (p_i - p_j) \cdot \mathrm{sim}(\mathrm{e_i}, \mathrm{e_j}) < \mathrm{C}$$

Putting it all together

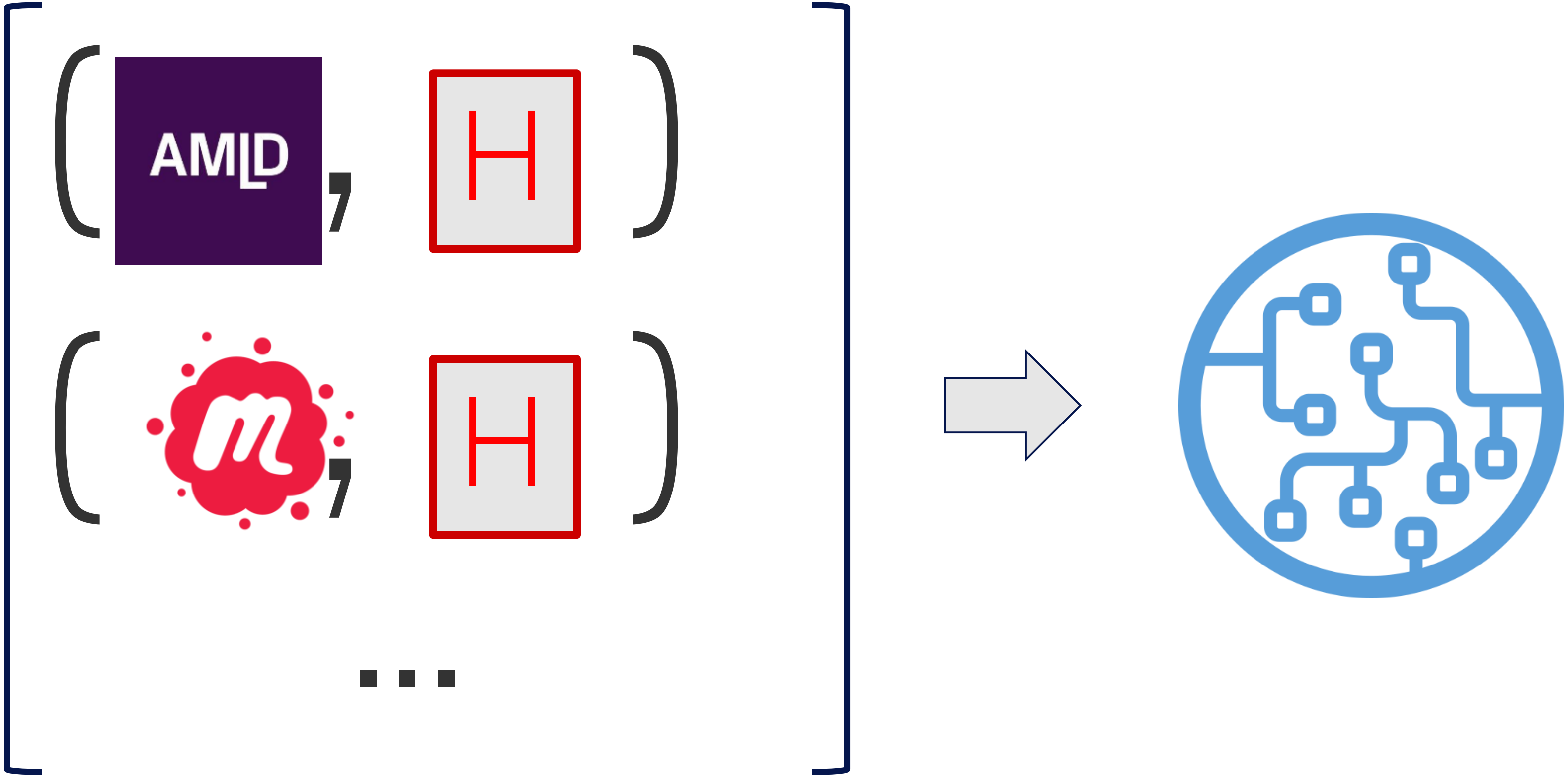
Maximize

$$P(\text{data}) = \prod_{D \in \mathcal{D}} P(D) \prod_{\overline{D} \in \overline{\mathcal{D}}} P(\overline{D}) \qquad \text{where } P(\overline{D}) = \prod_{e_i \in \overline{D}} (1 - p_{e_i})$$

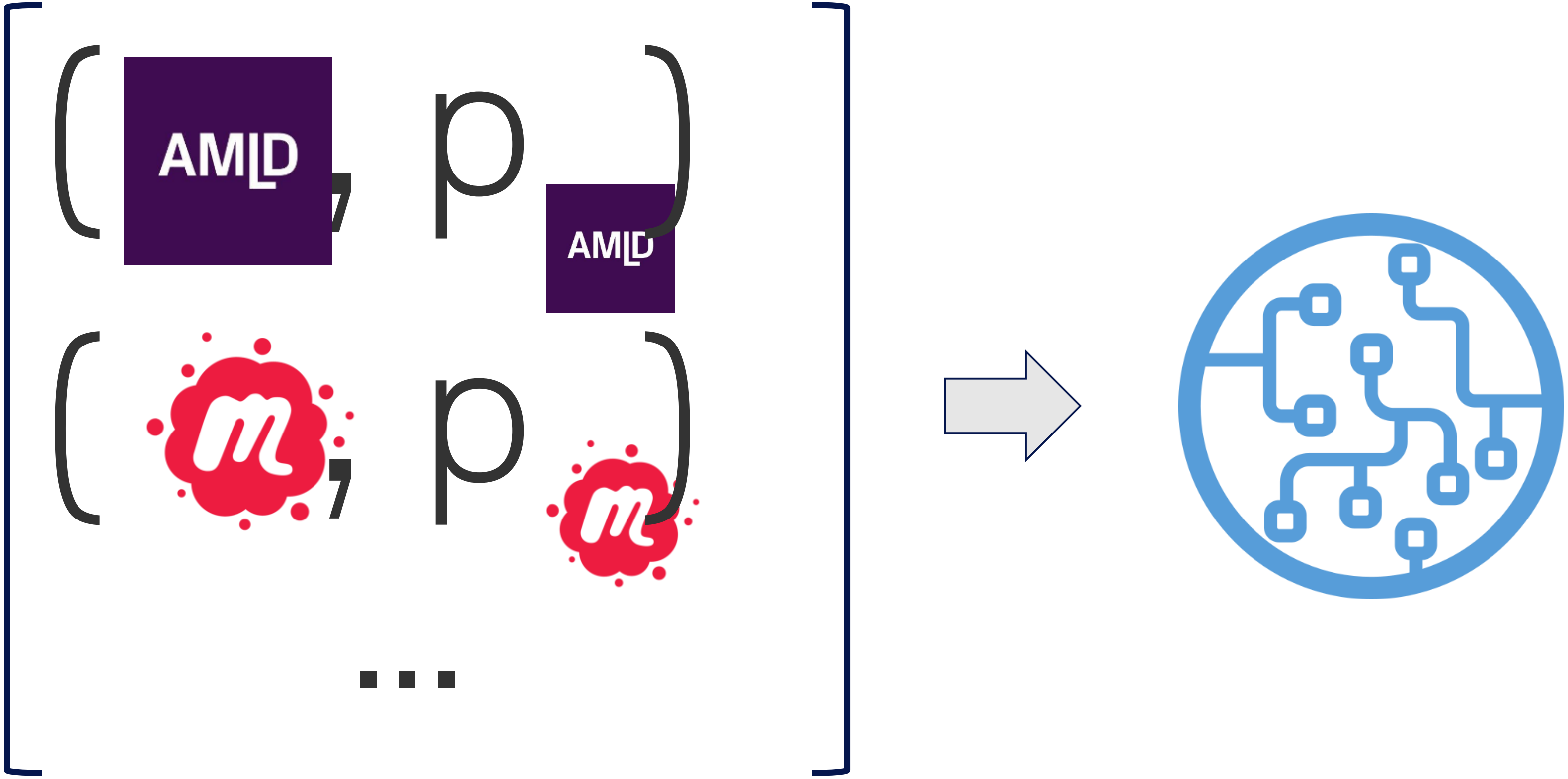
$$P(D) = 1 - \prod_{e_i \in D} (1 - p_{e_i})$$

$$\text{subject to } \sum_{i,j} (p_i - p_j) \cdot \mathrm{sim}(\mathrm{e_i}, \mathrm{e_j}) < \mathrm{C}$$

So instead of this...



… we can have this!



… we can have this!

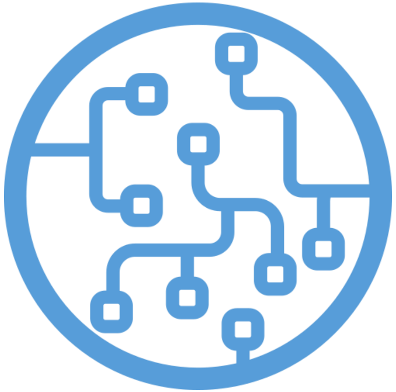
[

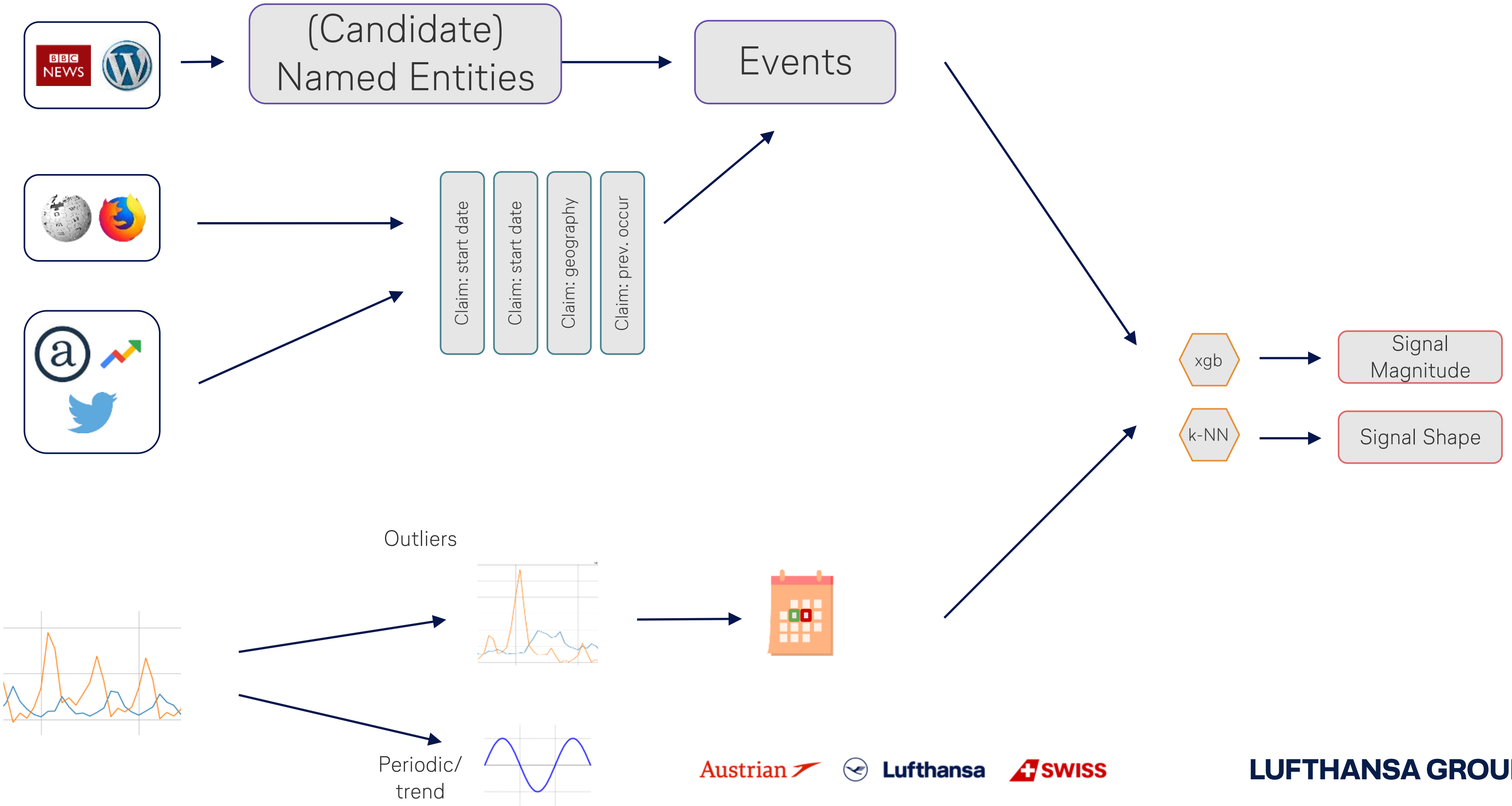
(AMLD, 0.9)

(, 0.1)

...

]




BBC NEWS
(Candidate)
Named Entities
Events
Claim: start date
Claim: start date
Claim: geography
Claim: prev. occur
xgb
k-NN
Signal
Magnitude
Signal Shape
Outliers
Periodic/
trend

Results: Event Information reveals significant Forecasting Improvement

Current Forecast

Austrian Lufthansa SWISS

LUFTHANSA GROUP

Results: Event Information reveals significant Forecasting Improvement

Current Forecast

+

Data Signal

Contextual Data

Triennial trade fair in **Stuttgart**
Categorization: **Expo**
Estimated attendance: **65,603**
Online + News presence: **medium**
Social media presence: **medium**

Results: Event Information reveals significant Forecasting Improvement

Current Forecast

+

Data Signal

=

Improved Forecast

Contextual Data

R+T

Triennial trade fair in **Stuttgart**
Categorization: **Expo**
Estimated attendance: **65,603**
Online + News presence: **medium**
Social media presence: **medium**

KPI

RMSE: -4.3%

Events: +792 detected

Outlook: Prove potential for Catching Signals in price elasticity estimation, vision to use additional data sources in all models and further departments

Done: Demand

Model: Demand Volume Forecast

Timeline: 3-month test case up to MAR 2019

Scope: 14 routes, 1.5 years training and 3 months validation data

Evaluation: Forecast vs. observed demand (RMSE)

Result: Significantly lower forecasting error of 4.3%

Limitation: No technical implementation so far

Ongoing: Price Elasticity

Model: Demand Volume Forecast

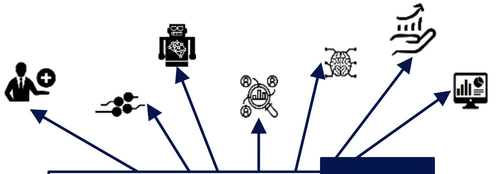
Timeline: 3-month test case up to APR 2020 incl. 6+ weeks of testing

Scope: 6+ ODs, 2 years training and 6+ weeks of validation data

Evaluation: Observed revenue during live test

Result: Clear potential detected in simulation, test evaluation awaited

Limitation: No proven revenue increase so far



Use enhanced analytical insights through enriching forecast models with contextual data

Vision: All Models

- Revenue Management
- Sales
- Marketing
- ...


Thank you
for your attention