

A Connected sky: forecasting the routes of the future

Network Science Project

Enrico Gottardis

11 September 2024

Table of contents



Introduction



Analysis and study of the networks

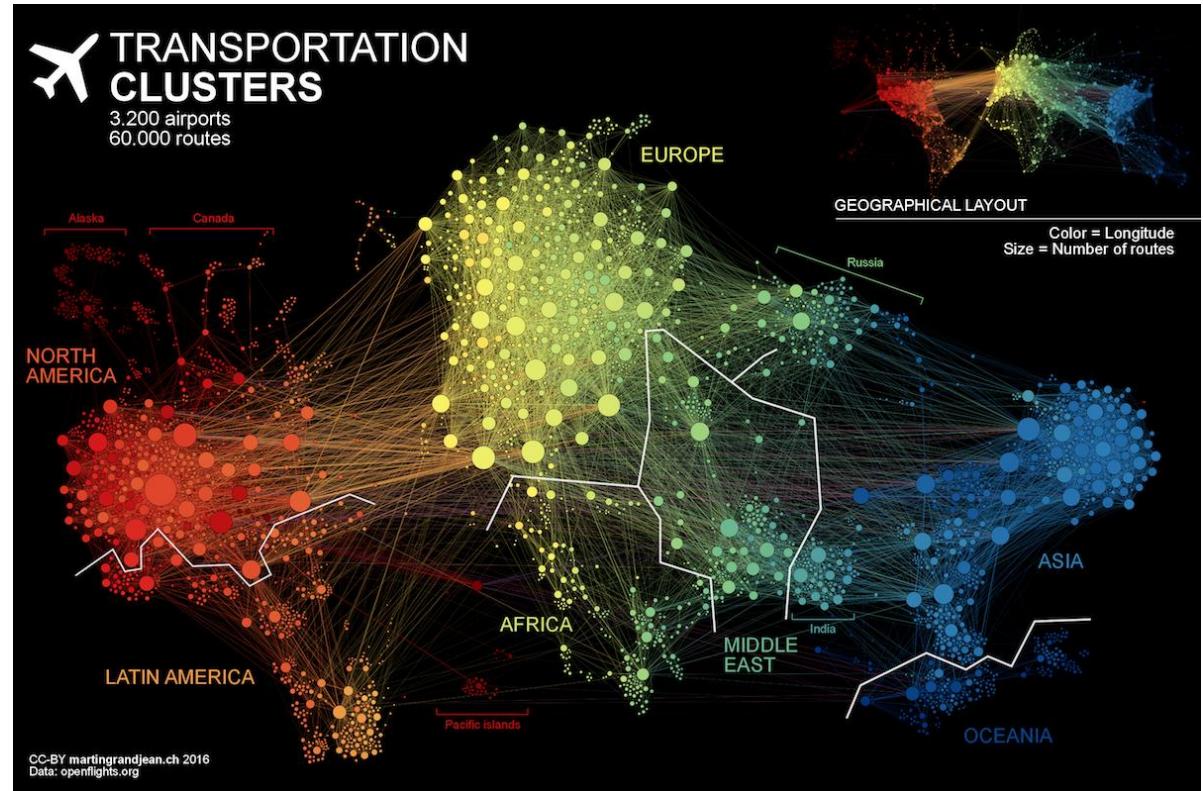


Link Prediction



Conclusion and final thoughts

Introduction - Inspiring works



War impact on russian airways – Powerpoint presentation by Zamengo Fulvio



MARTIN GRANDJEAN



<https://www.martingrandjean.ch/connected-world-air-traffic-network/>

Introduction - goals



Deep insights of the
analyzed networks



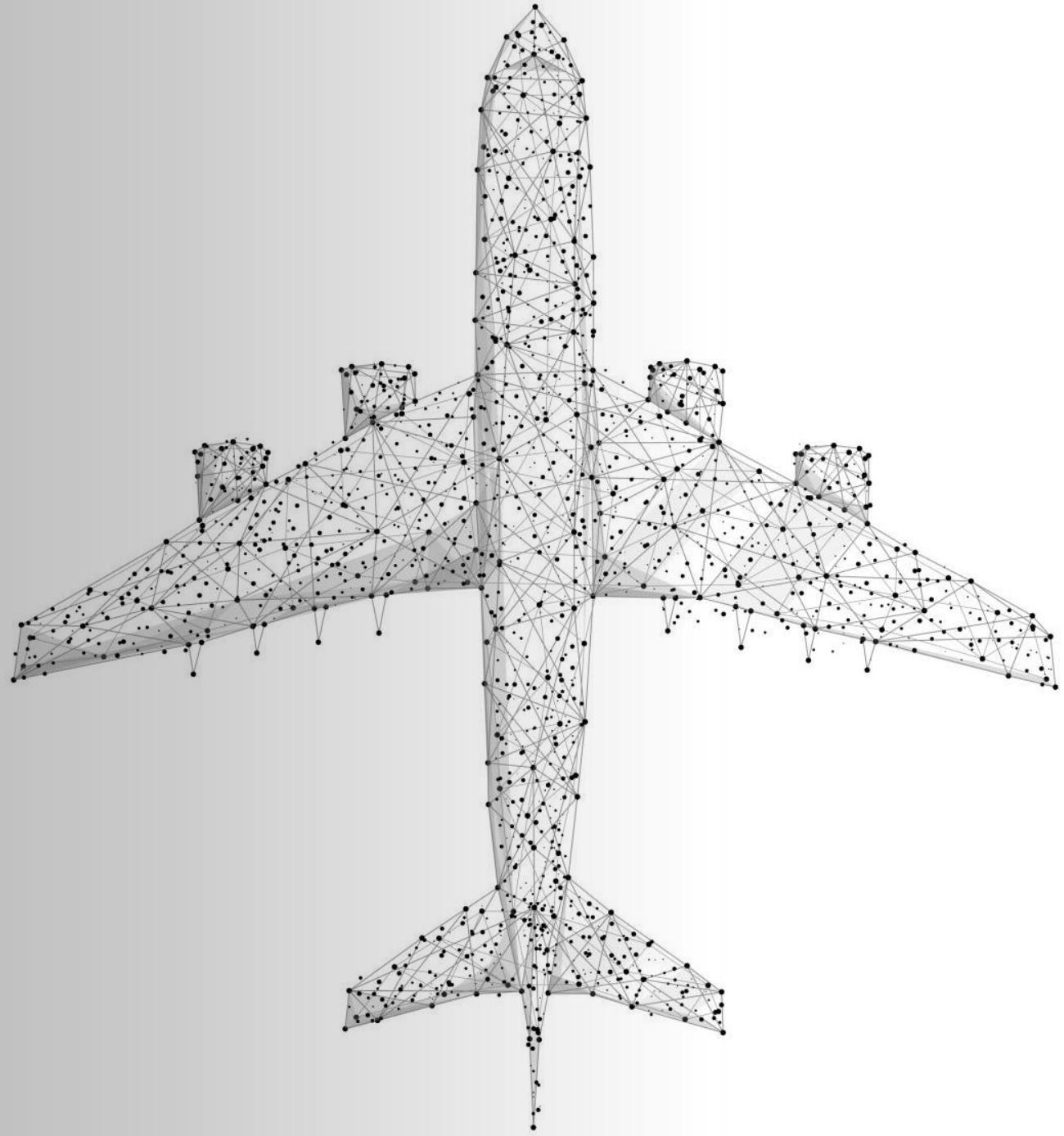
Prediction of new routes
based on the graphs studied



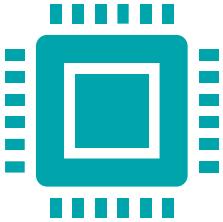
Verification whether the
predictions have been
accurate

Analysis and study of the network

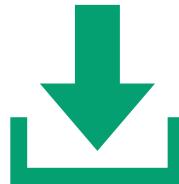
- Data description
- Centrality
- Community detection



Data description - Source



Data has been offered by the [OpenSky Network](#) website, which is well-known in the field of data analysis as regards sky traffic.



Data is free to download, though an account (which is free to create) is requested in order to do so.



The dataset considered is packed with information, much of which has been scraped in the preprocessing process. Furthermore, as the number of flights in a year is huge, the analysis is focused on the period that precedes the summer, in particular the study spans over the months [April, May and June](#) of the years [2020, 2021 and 2022](#).

Data description - Dataset (Flights)

The dataset considered is divided based on the year. An example of the polished dataset is given at the bottom of the page.

After scraping off the extra information, the dataset shows (for the span of time considered):

- 2M+ flights in 2020, with 12769 active airports
- 4M+ flights in 2021, with 14159 Active airports
- Almost 6M flights in 2022, with 14676 Active airports

Important notice: rows in the dataset containing missing data have been removed from the polished dataset

Origin	Destination	Origin index	Destination index
KLAX	EHAM	8271	5674

Data description - Dataset (Airports)

Information regarding the geographical location of the airports is found in the same dataset containing the flights. However, in order to separate tasks and better manage information, I decided to split information regarding flights and that regarding airports on their own.

The airports are expressed using the ICAO airport code, a 4-letter code, different than the well-known 3-letter IATA codes.

Here is an example of how the polished data about airports is presented:

Origin	Latitude	Longitude
KLAX	33.935930	-118.418083

Insight of the graphs

2020

- Nodes (Airports): 12769
- Edges (Air routes): 321686
- Flights: 2253052
- Density: 0.299%
- Giant strong cc size: 12725

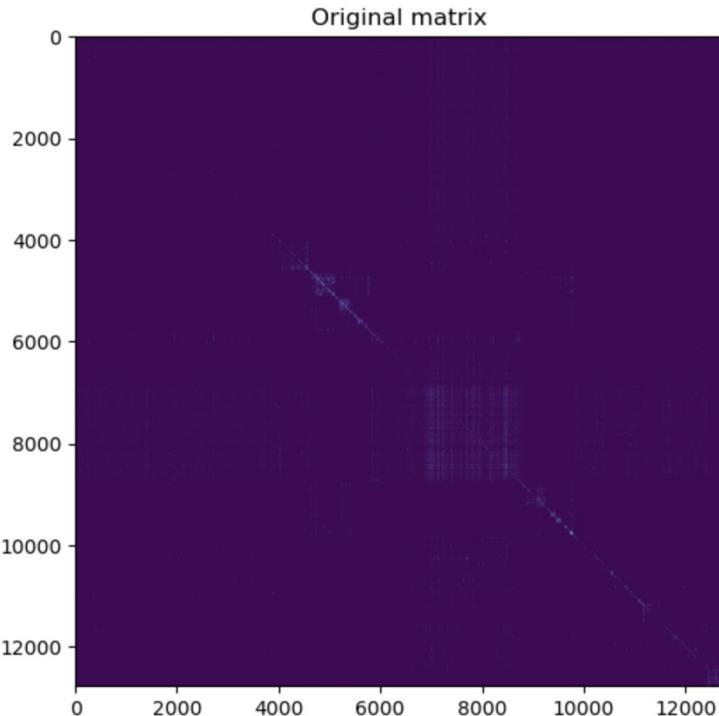
2021

- Nodes (Airports) : 14159
- Edges (Air routes): 490410
- Flights : 4588501
- Density: 0.367%
- Giant strong cc size: 14082

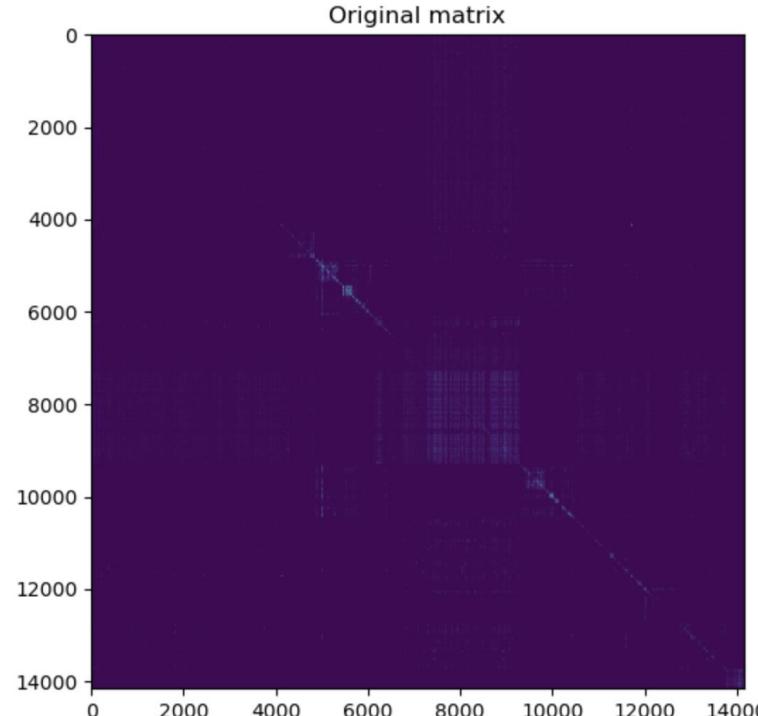
2022

- Nodes (Airports): 14676
- Edges (Air routes): 544484
- Flights : 5894448
- Density: 0.382%
- Giant strong cc size: 14644

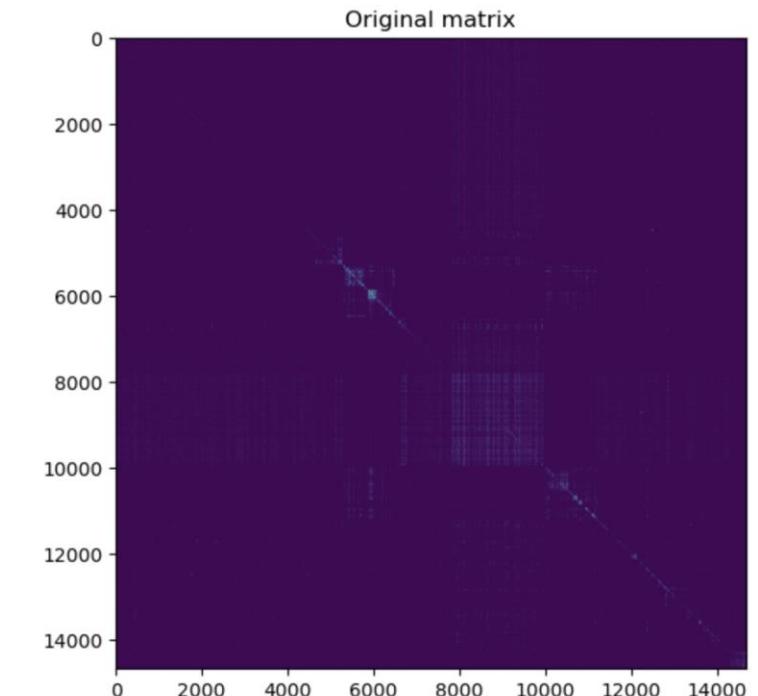
Data description - Adjacency matrix



2020



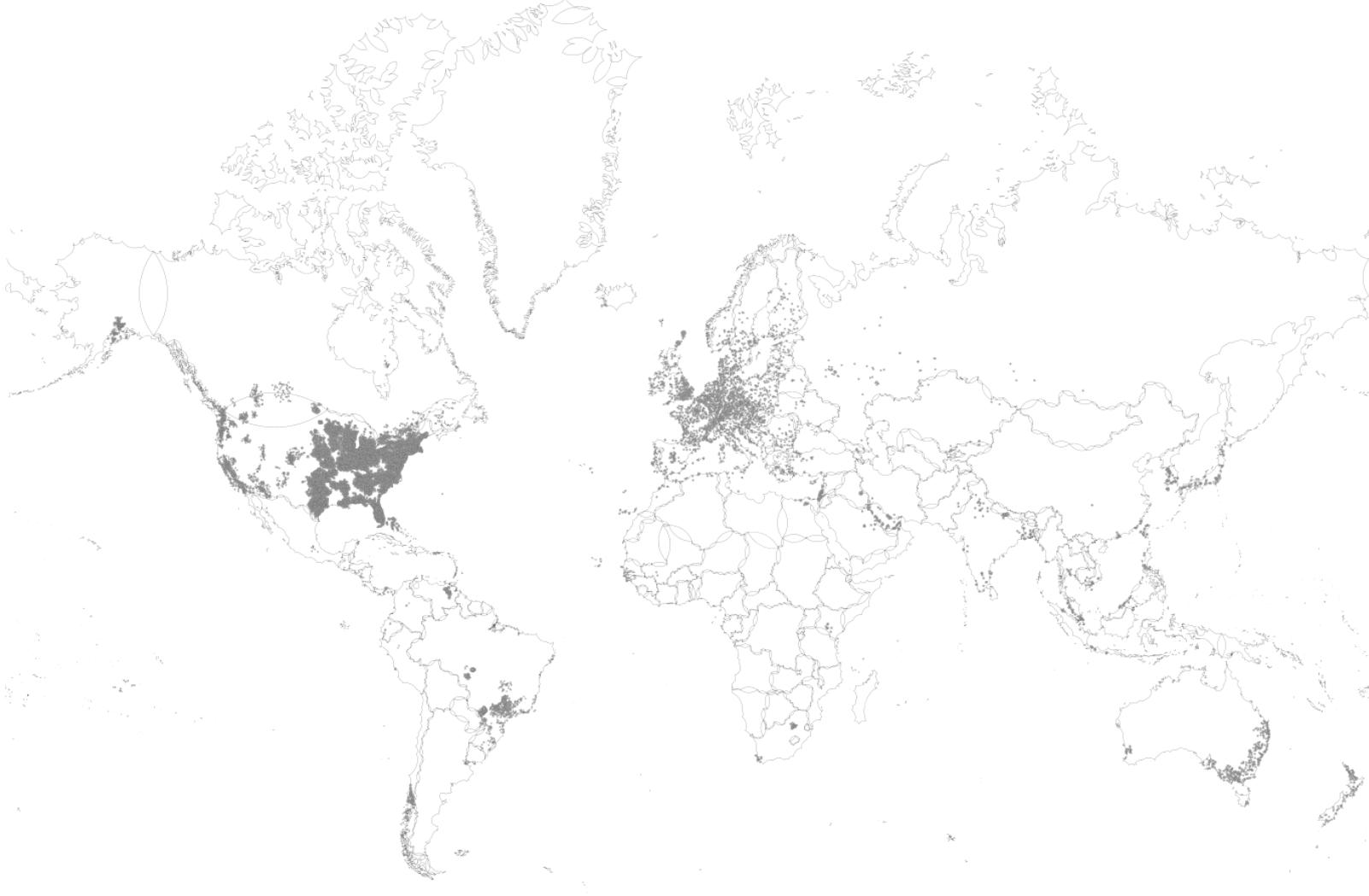
2021



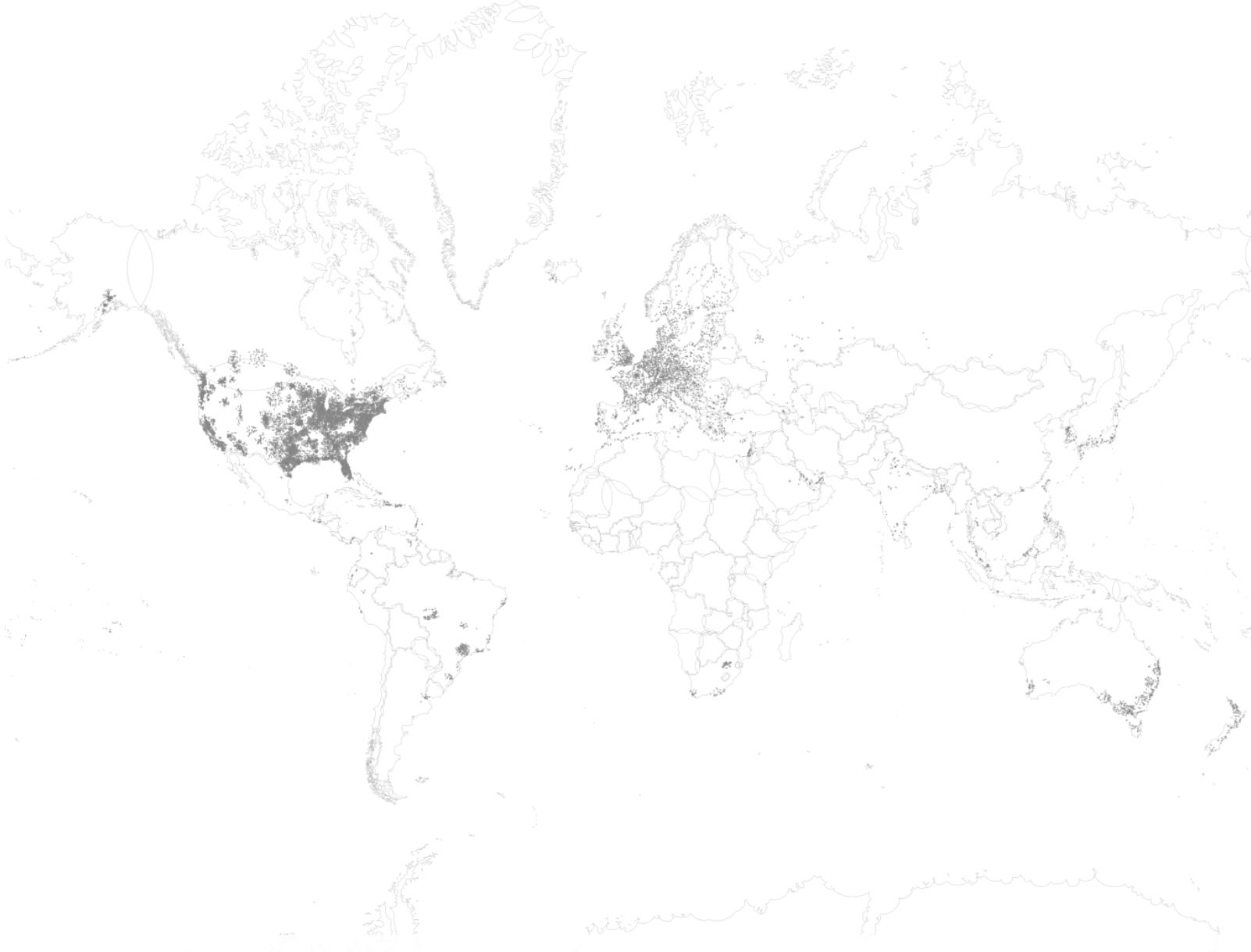
2022



Active airports 2020



Active airports 2021



Active airports 2022



Flight routes 2020



Flight routes 2021



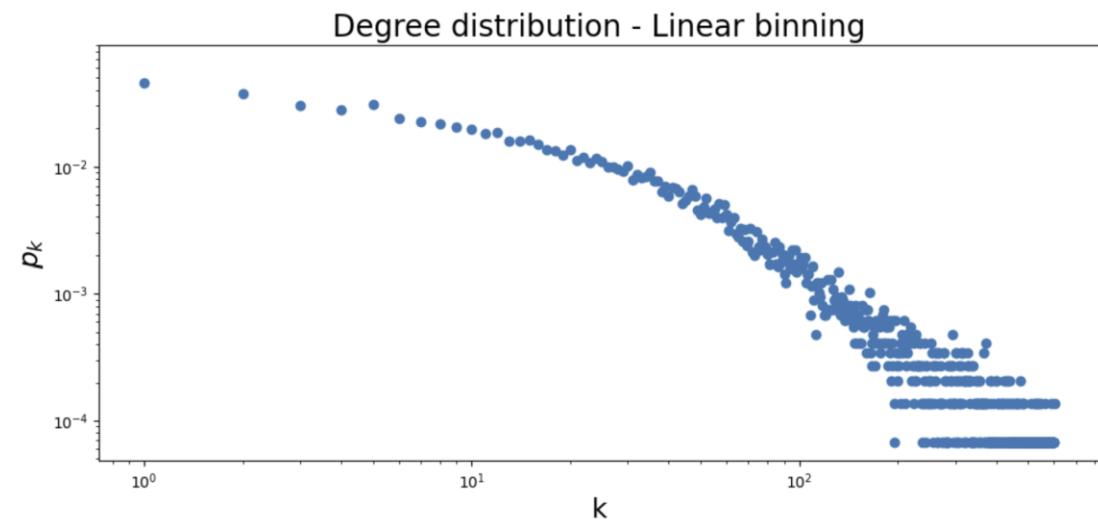
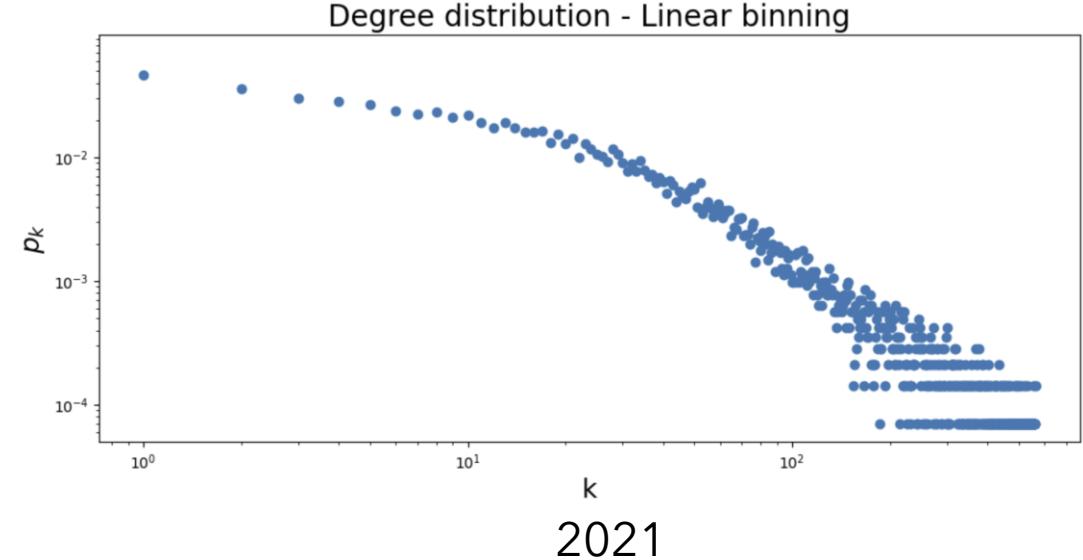
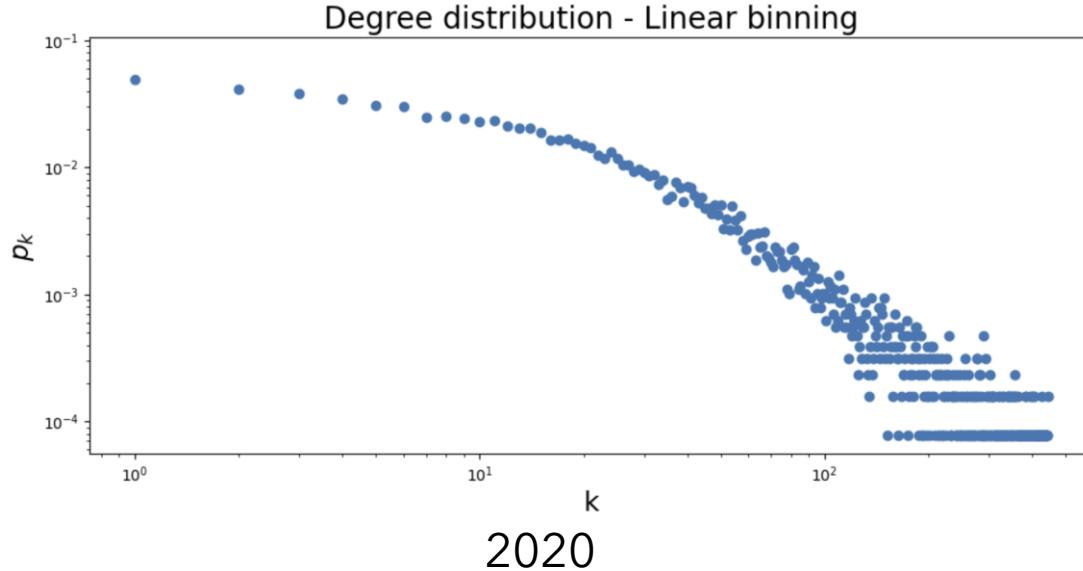
Flight routes 2022

Centrality measures:

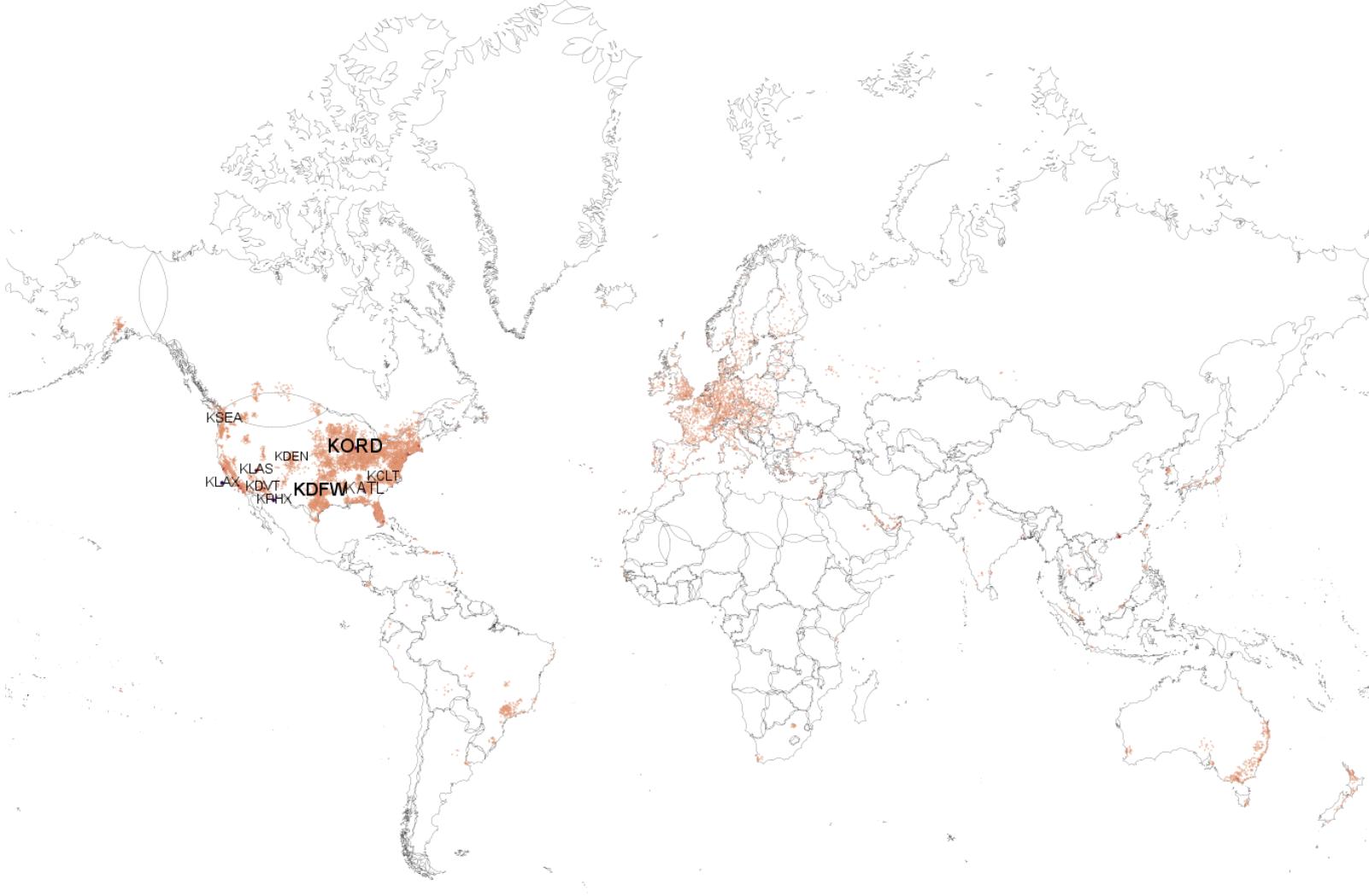
- Degree
- HITS (Authority + Hubs)
- PageRank



Degree distributions of the airports



All years share the same shape of distribution: it follows the power law, with many low degrees and some hubs.



Airports 2020 – Relevance by degree

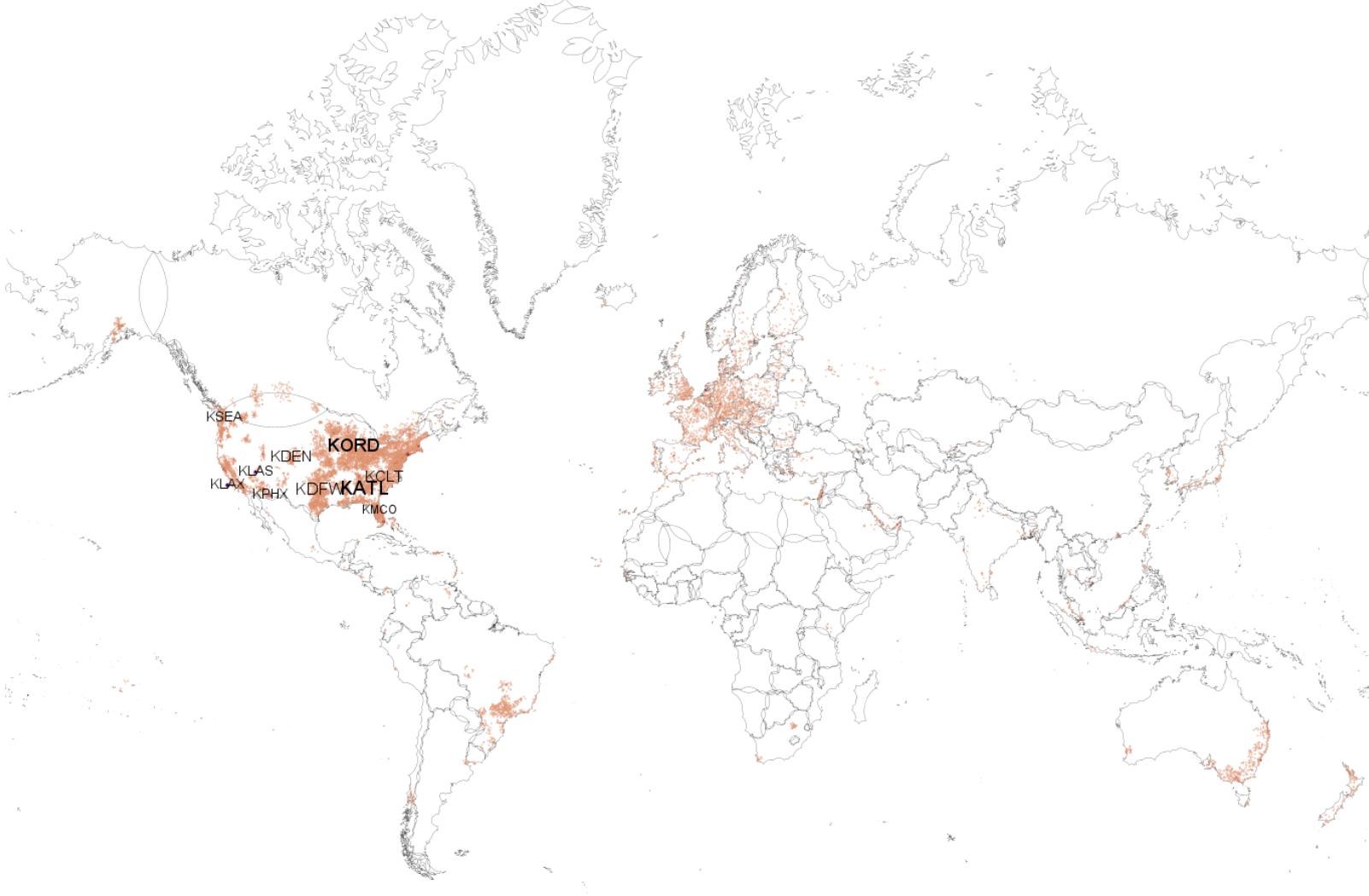
Airports 2020 - Degree centrality

Most important airports
with weights set to 1:

KORD	Chicago-O'Hare
KDFW	Dallas-Fort Worth
KCLT	Charlotte-Douglas
KAPA	Centennial airport, Colorado
KDAL	Dallas-Love
KADS	Addison, Dallas
KIAD	Washington-Dulles
KPDK	DeKalb-Peachtree, Atlanta
KPWK	Chicago Executive Airport
KAPF	Naples airport, Florida

Most important airports
with right weights:

KORD	Chicago-O'Hare
KDFW	Dallas-Fort Worth
KATL	Hartsfield-Jackson, Atlanta
KPHX	Phoenix-Sky
KLAX	Los Angeles Airport
KCLT	Charlotte-Douglas
KSEA	Seattle-Tacoma International Airport
KDVT	Phoenix Deer Valley
KLAS	Las Vegas Harry Reid
KDEN	Denver International Airport



Airports 2021 – Relevance by degree

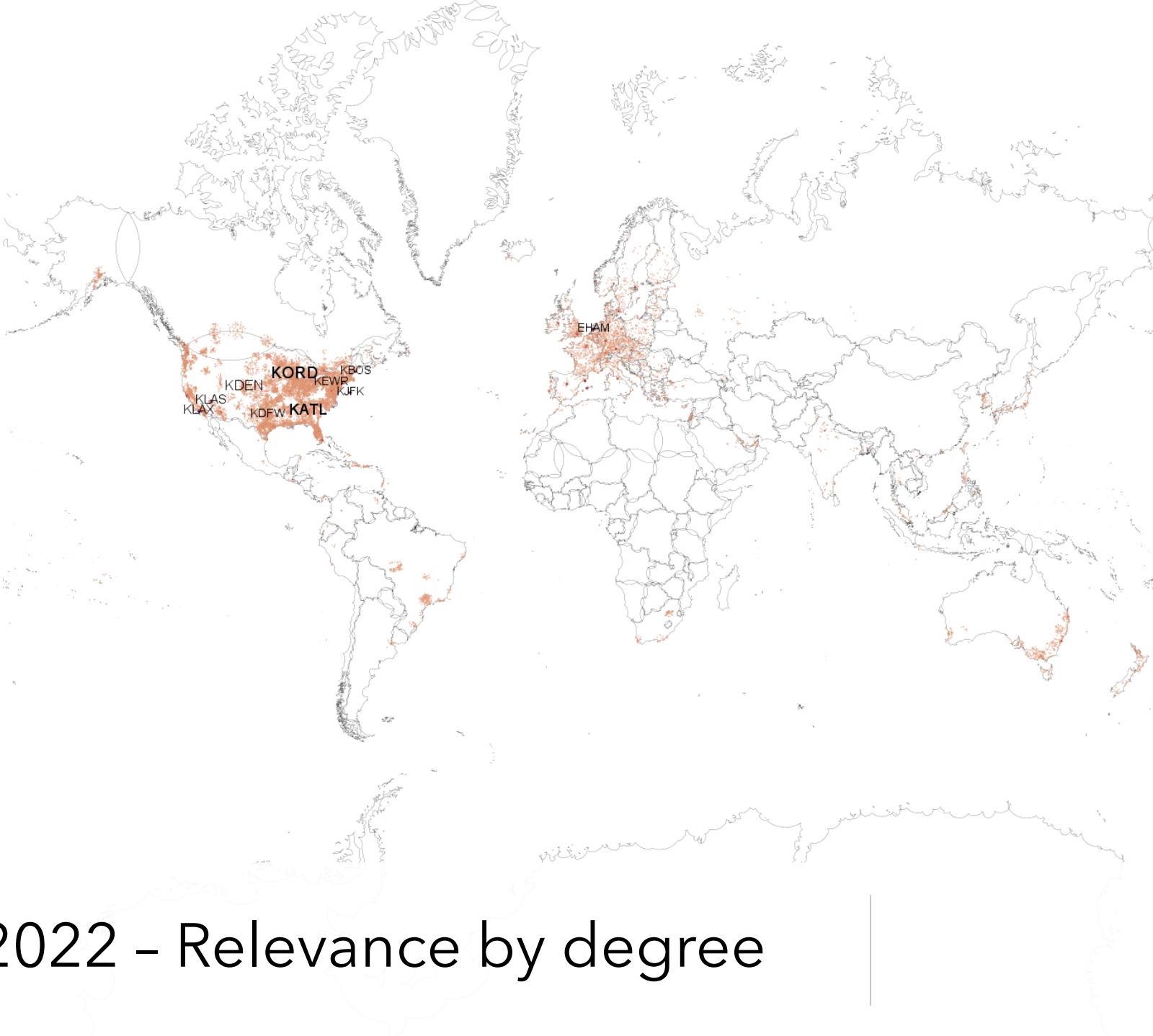
Airports 2021 - Degree centrality

Most important airports
with weights set to 1:

KTEB	Teterboro Airport
KCLT	Charlotte-Douglas
KDAL	Dallas-Love
KPDK	DeKalb-Peachtree, Atlanta
KHPN	Westchester County Airport
	Nashville International Airport (Tennessee)
KBNA	
KIAD	Washington-Dulles
KORD	Chicago-O'Hare
KDFW	Dallas-Fort Worth
KPWK	Chicago Executive Airport

Most important airports
with right weights:

KATL	Hartsfield-Jackson Atlanta
KORD	Chicago-O'Hare
KDFW	Dallas-Fort Worth
KDEN	Denver International Airport
KCLT	Charlotte-Douglas
KLAX	Los Angeles Airport
KLAS	Las Vegas Airport
KPHX	Phoenix-Sky
KSEA	Seattle-Tacoma International Airport
KMCO	Orlando International Airport



Airports 2022 – Relevance by degree

Airports 2022 - Degree centrality

Most important airports
with weights set to 1:

KTEB	Teterboro Airport
KDAL	Dallas-Love
KIAD	Washington-Dulles
KBNA	Nashville International Airport (Tennessee)
KPDK	DeKalb-Peachtree Airport (Atlanta)
KHPN	Westchester County Airport
KDFW	Dallas-Fort Worth
KCLT	Charlotte-Douglas
KORD	Chicago-O'Hare
KPWK	Chicago Executive Airport

Most important airports
with right weights:

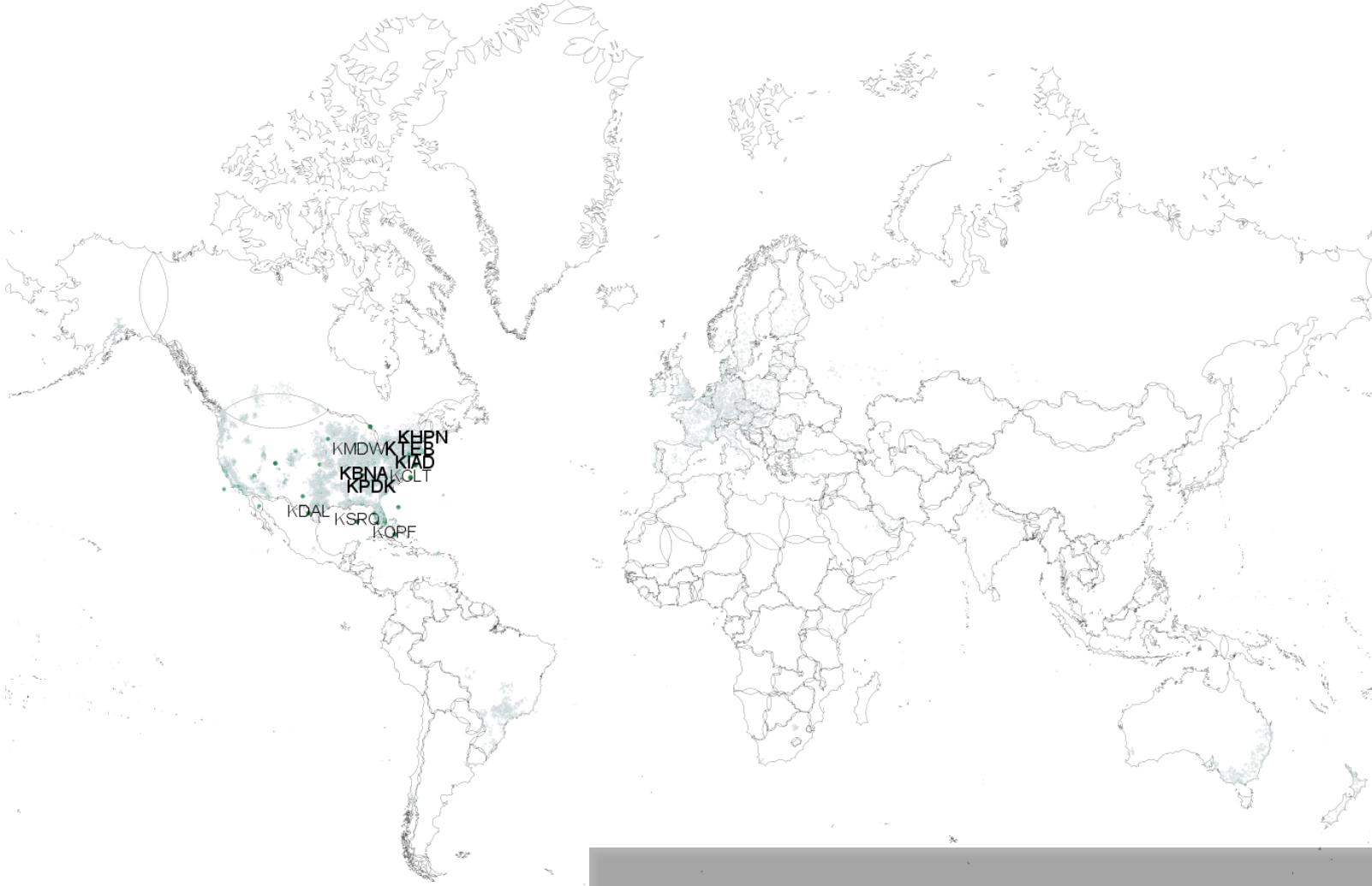
KORD	Chicago-O'Hare
KATL	Hartsfield-Jackson Atlanta
KDEN	Denver International Airport
KLAS	Las Vegas Airport
KLAX	Los Angeles Airport
KDFW	Dallas-Fort Worth
KEWR	Newark Liberty International
EHAM	Amsterdam Schiphol Airport
KBOS	Boston Logan International Airport
KJFK	JFK - New York



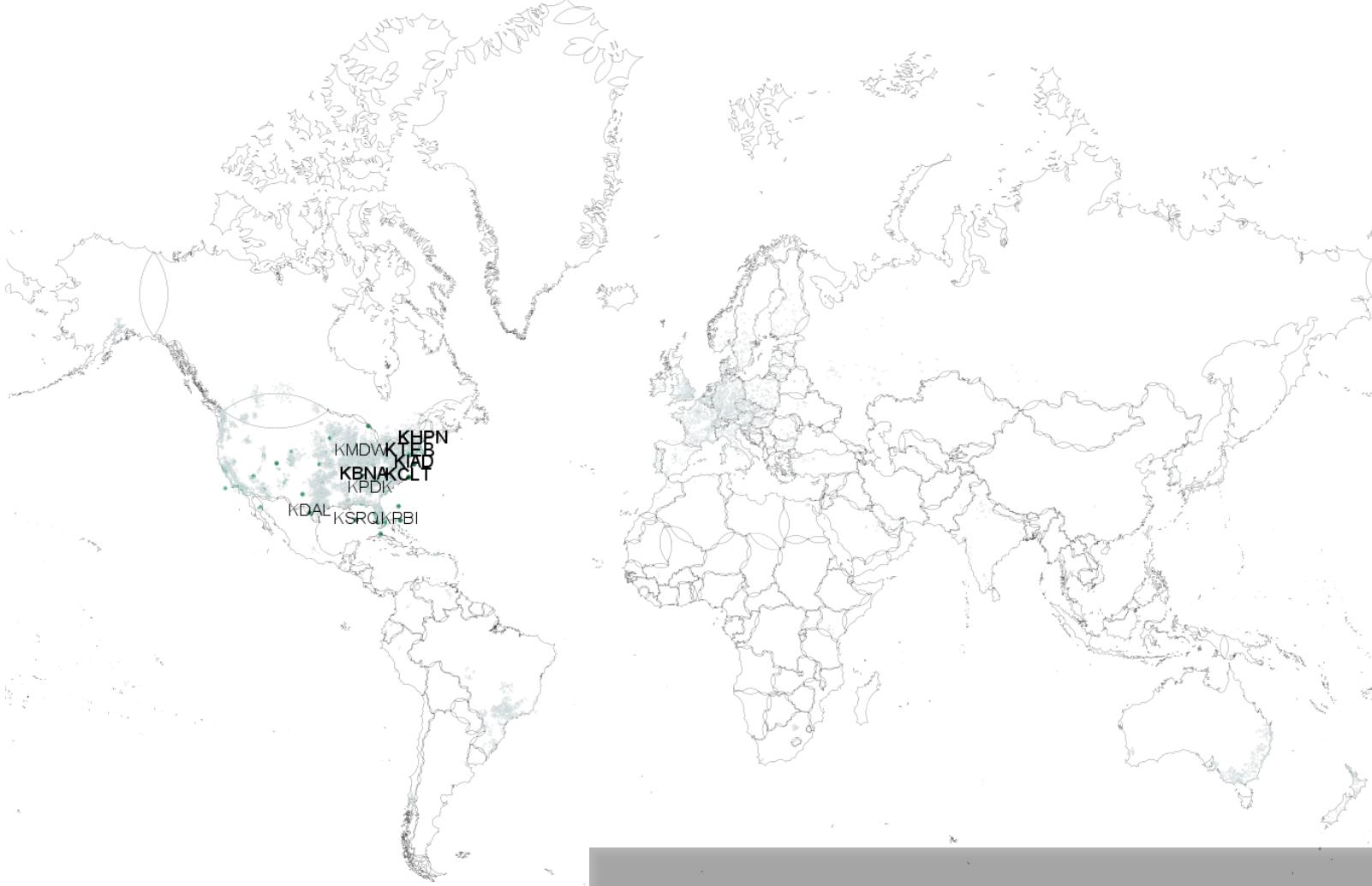
Airports 2020+ Authority



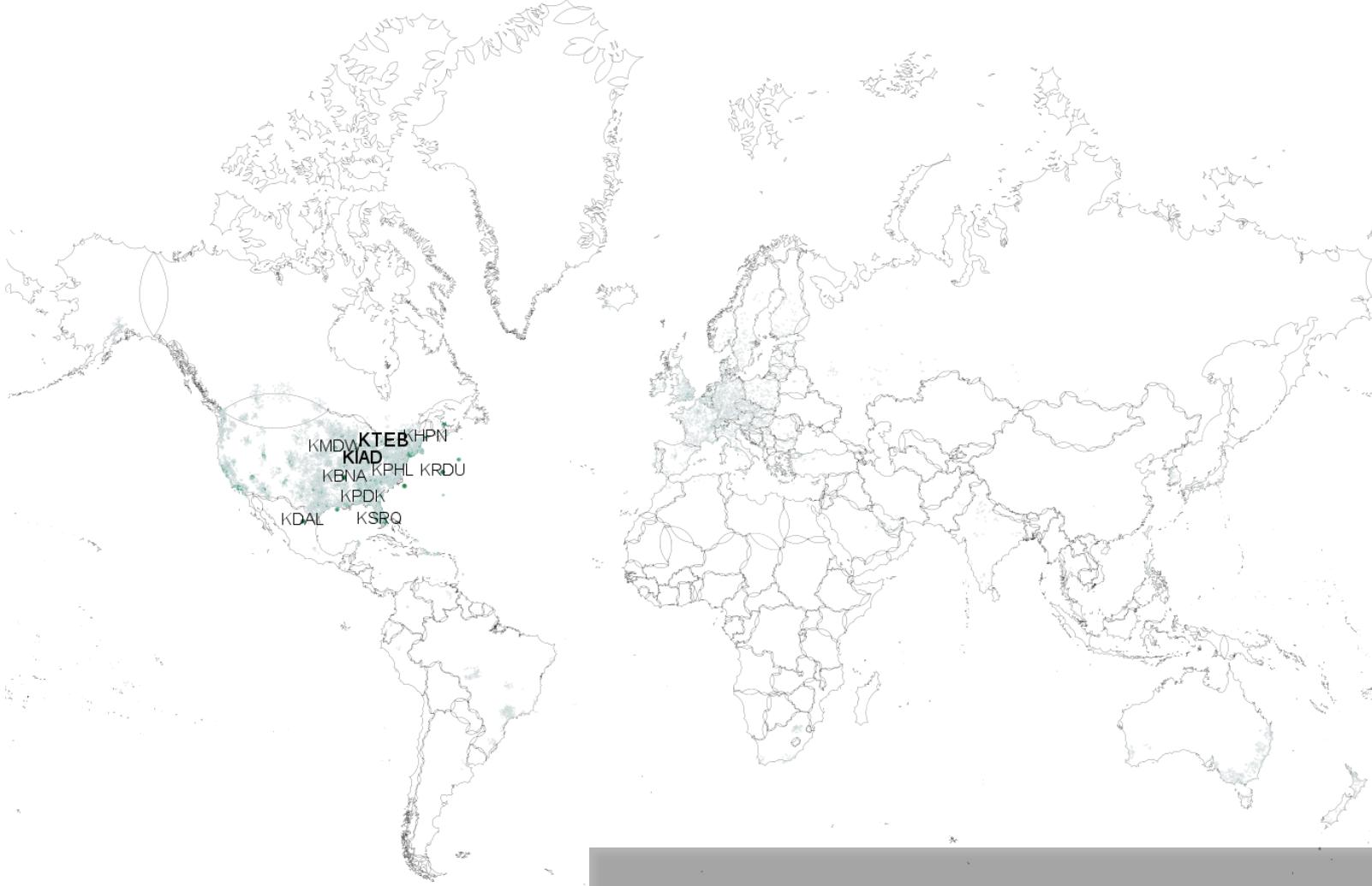
Airports 2020 -
Hubs



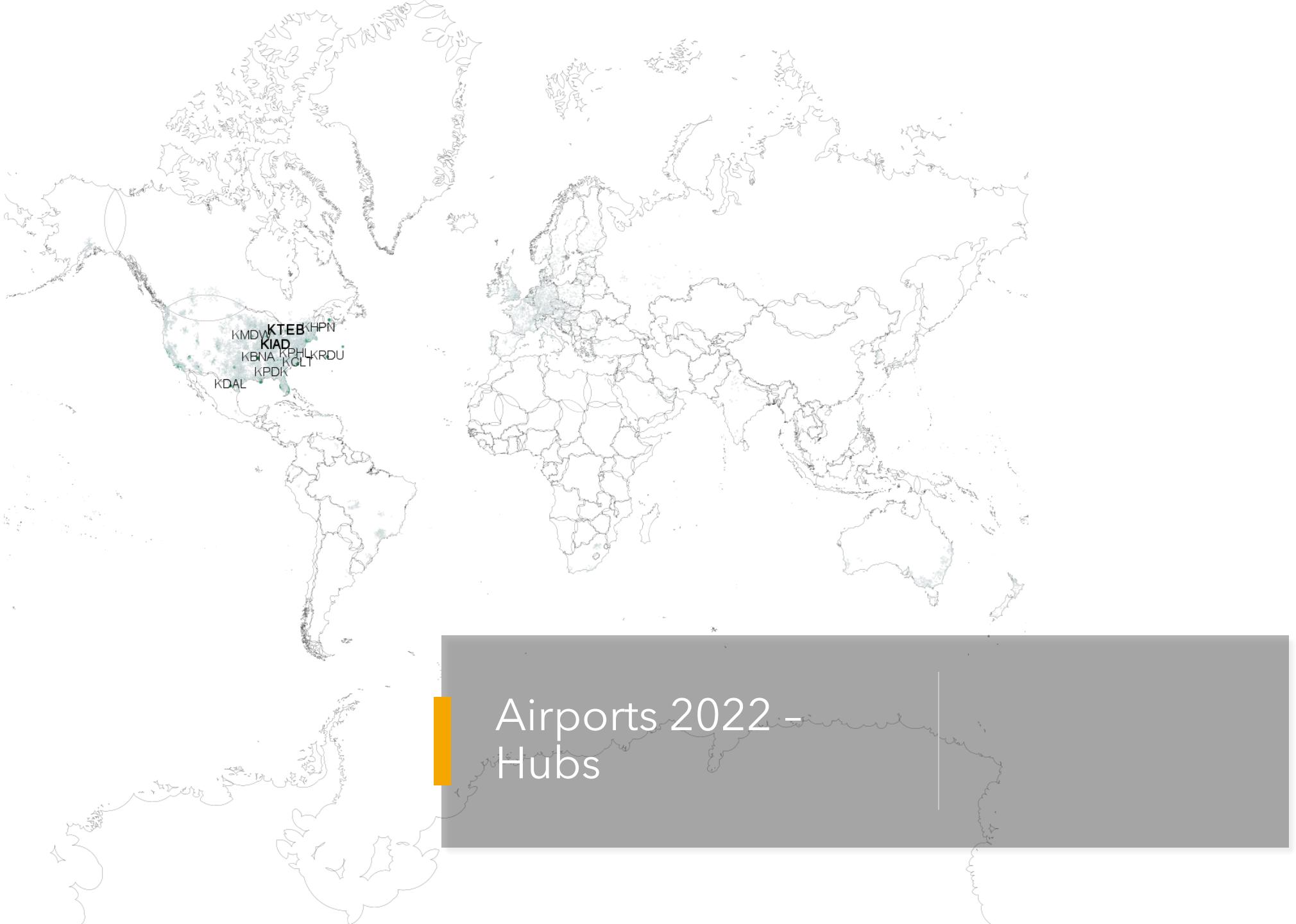
Airports 2021 -
Authority

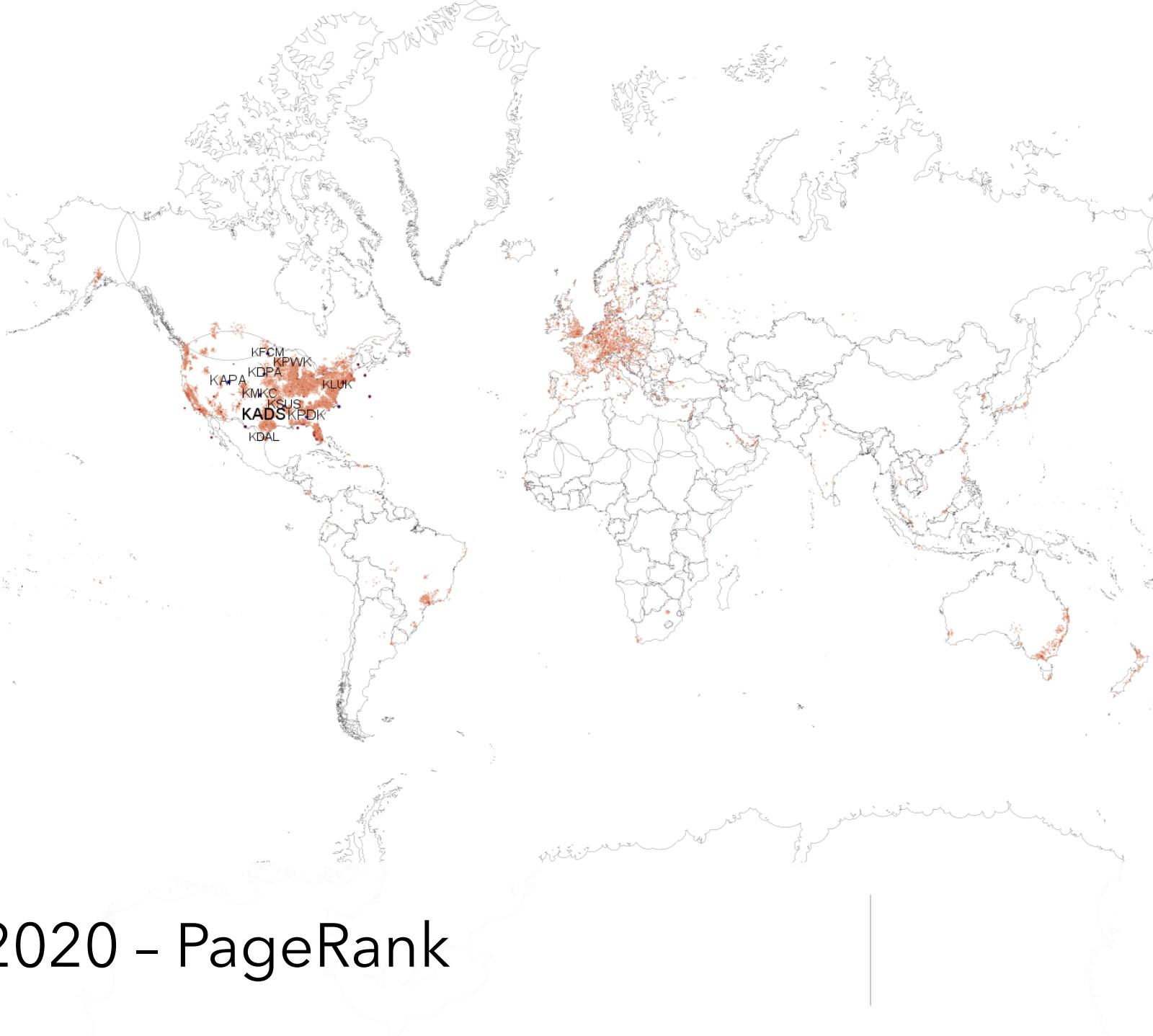


Airports 2021 -
Hubs

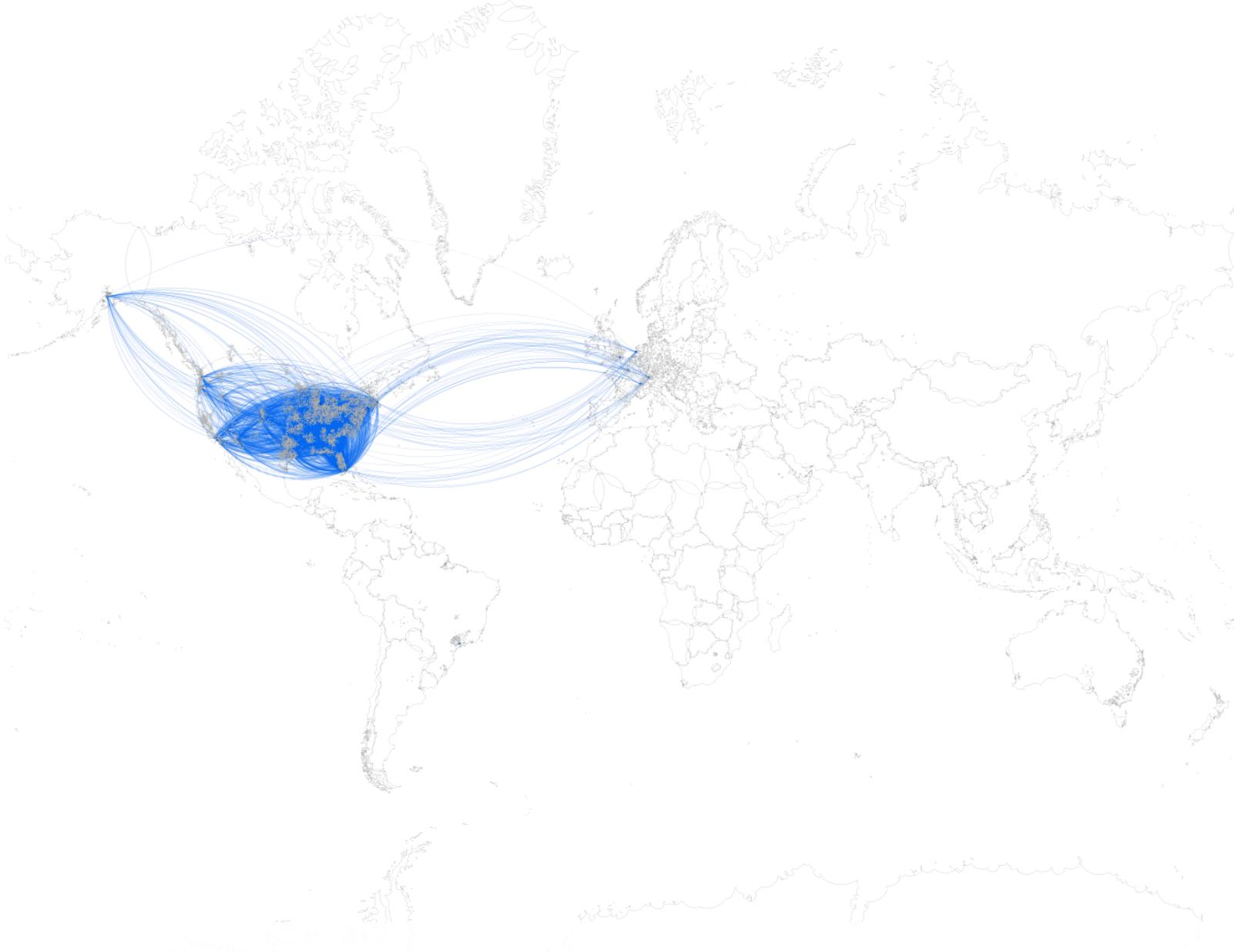


Airports 2022-
Authority

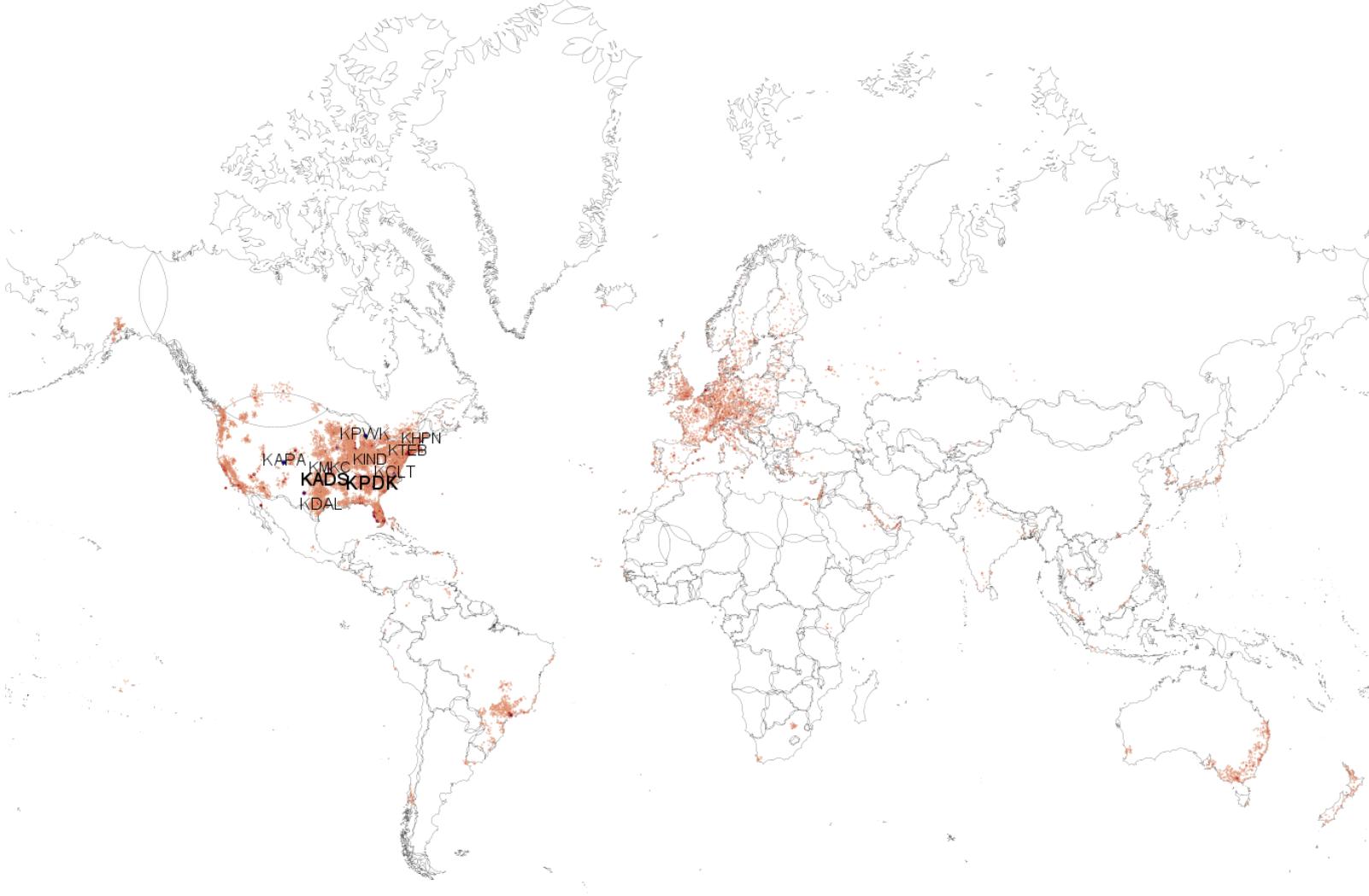




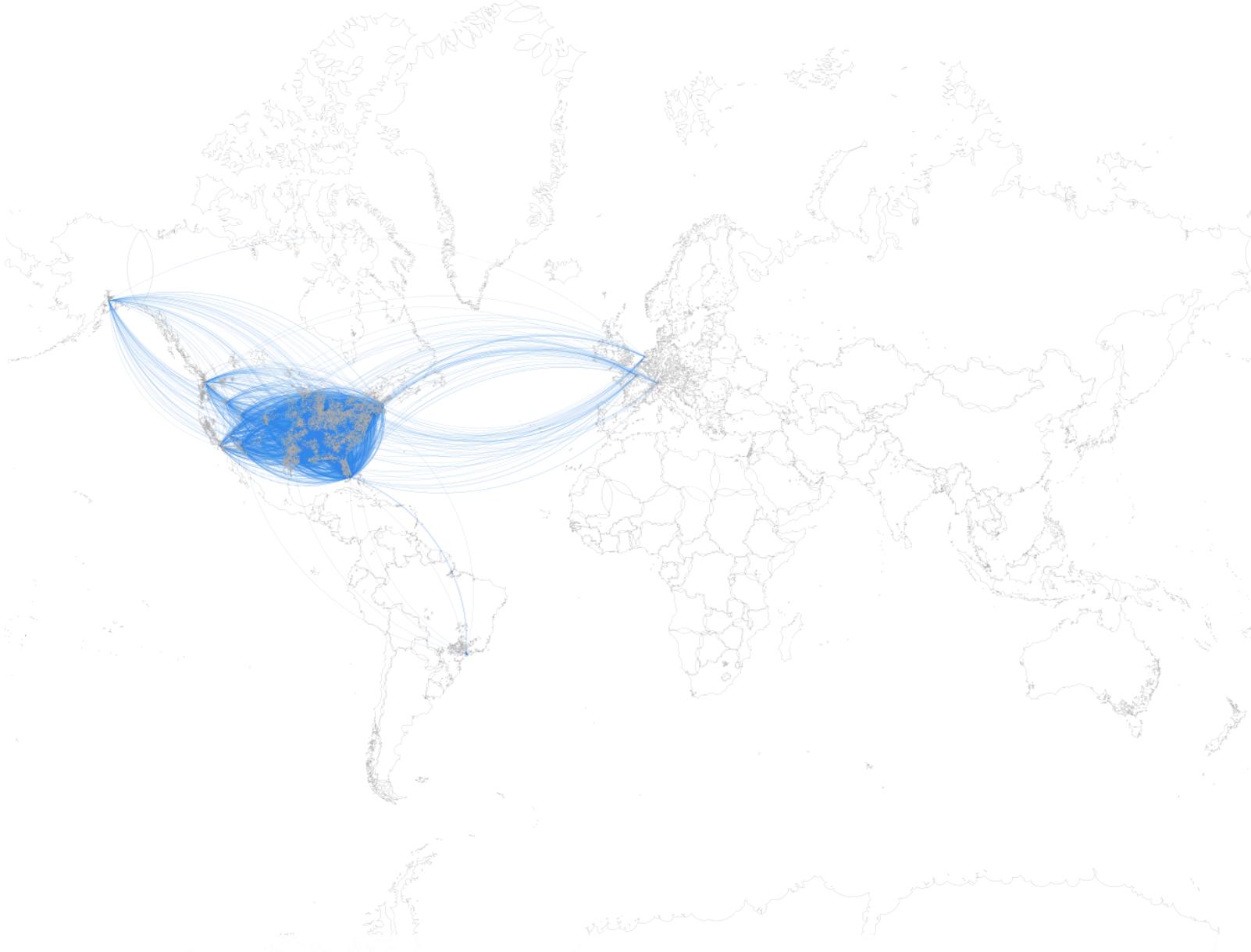
Airports 2020 - PageRank



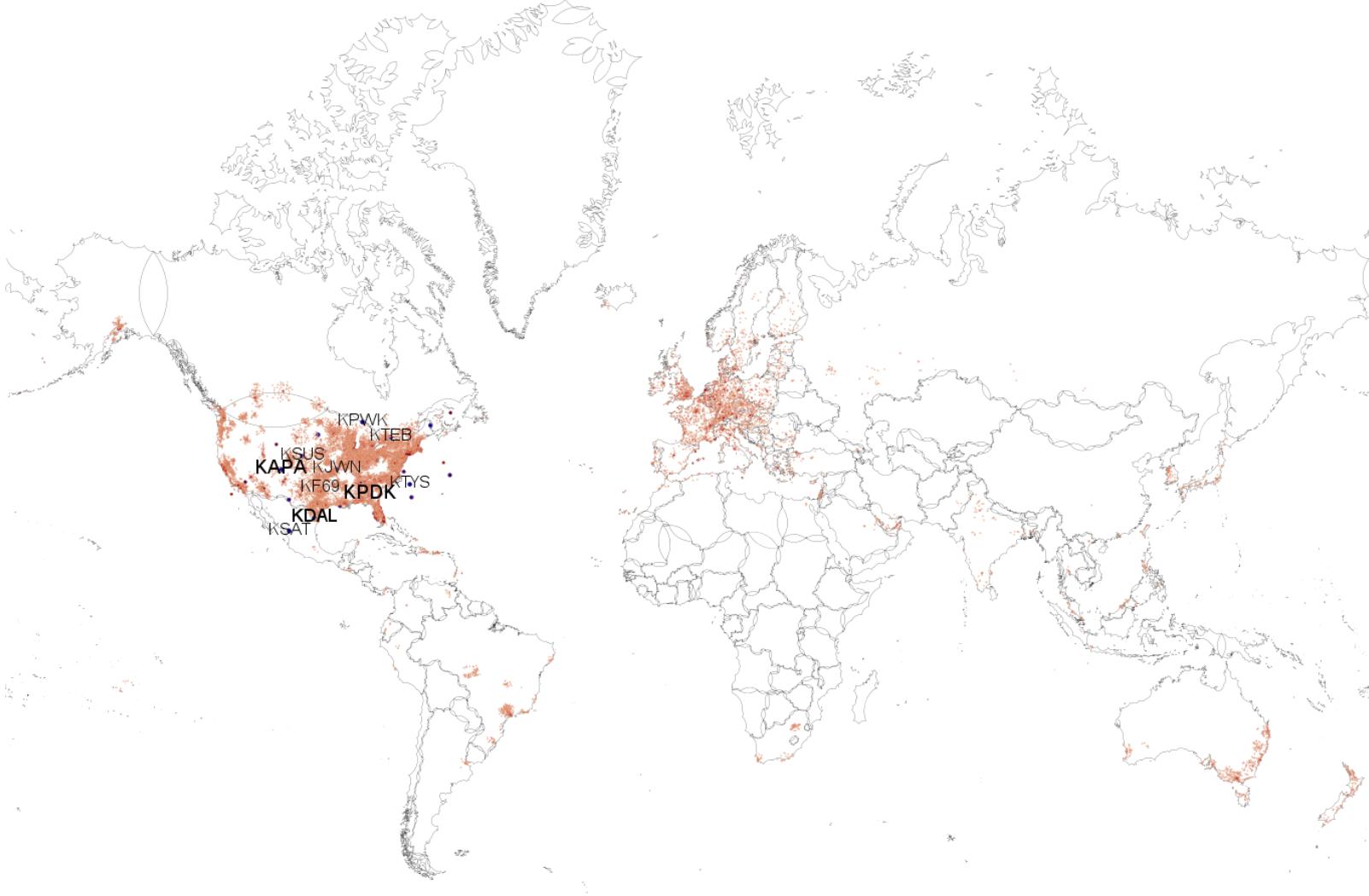
Top 1% Airports - 2020



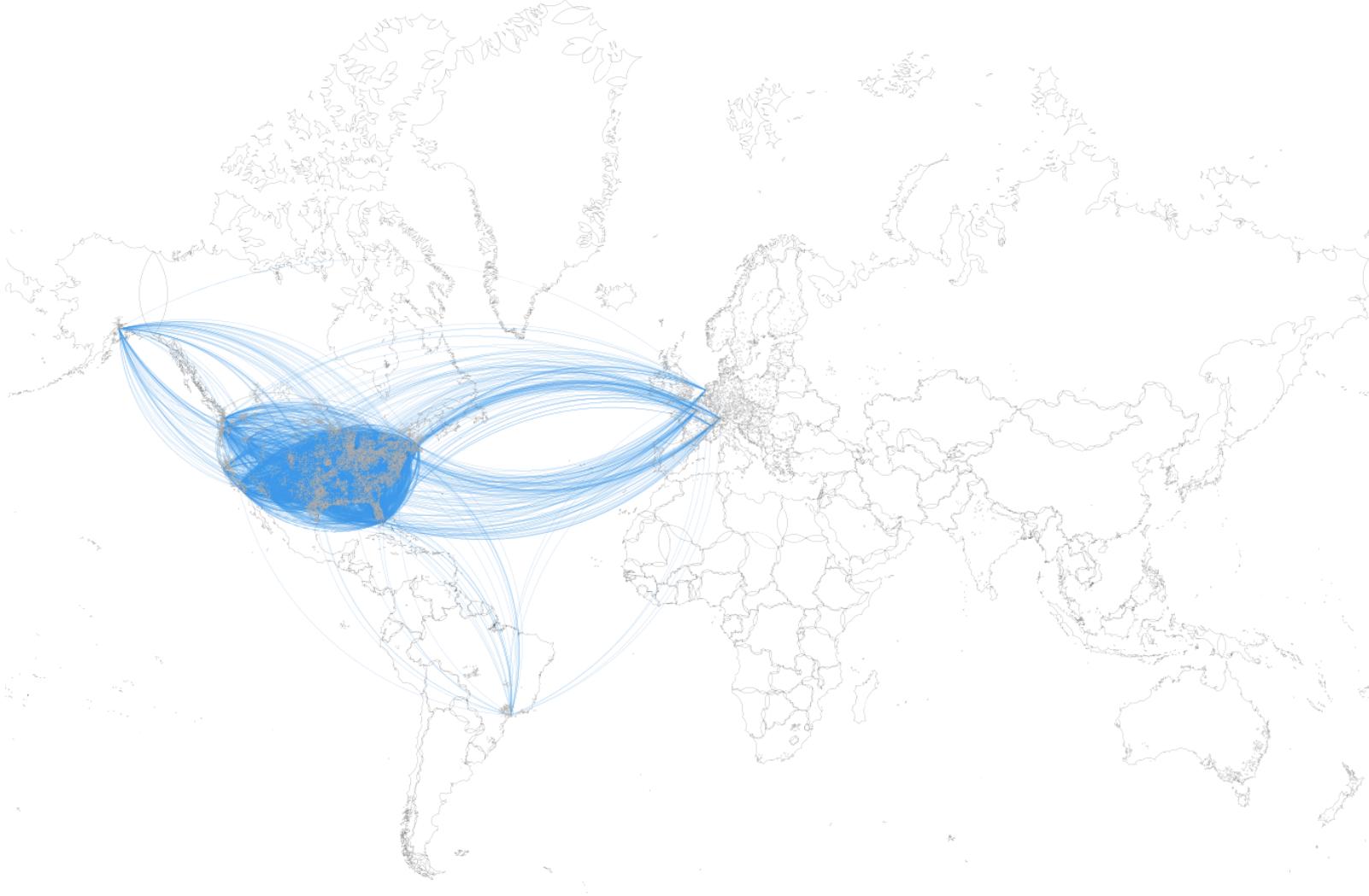
Airports 2021 - PageRank



Top 1% Airports - 2021



Airports 2022 - PageRank



Top 1% Airports - 2022

Comparing PageRank measures

2020

KADS
KAPA
KPDK
KPWK
KSUS
KDPA
KMKC
KDAL
KFCM
KLUK

2021

KPDK
KADS
KCLT
KDAL
KPWK
KAPA
KHPN
KTEB
KIND
KMKC

2022

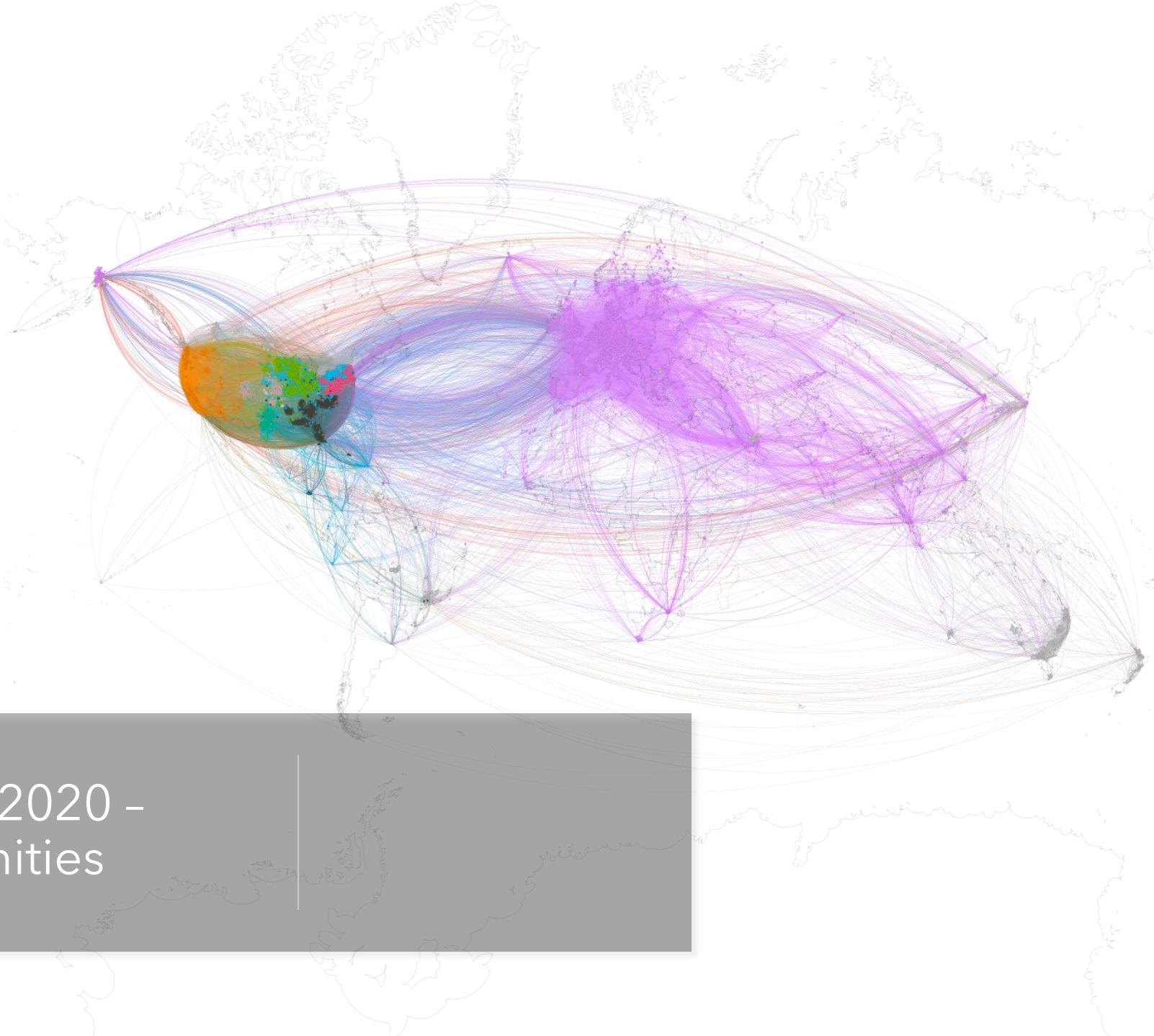
KPDK
KAPA
KDAL
KSAT
KPWK
KTEB
KJWN
KSUS
KF69
KTYS

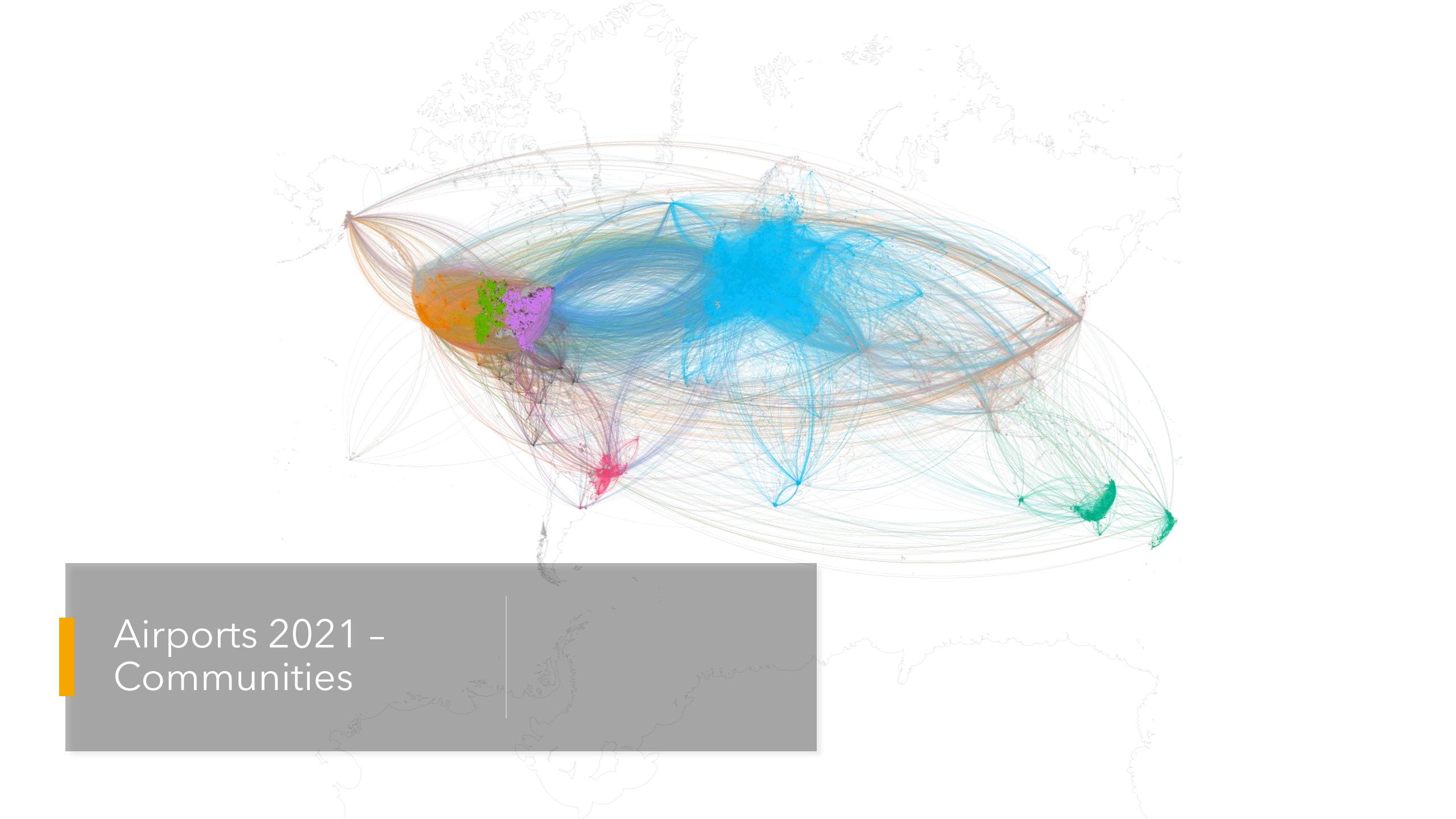


Community detection



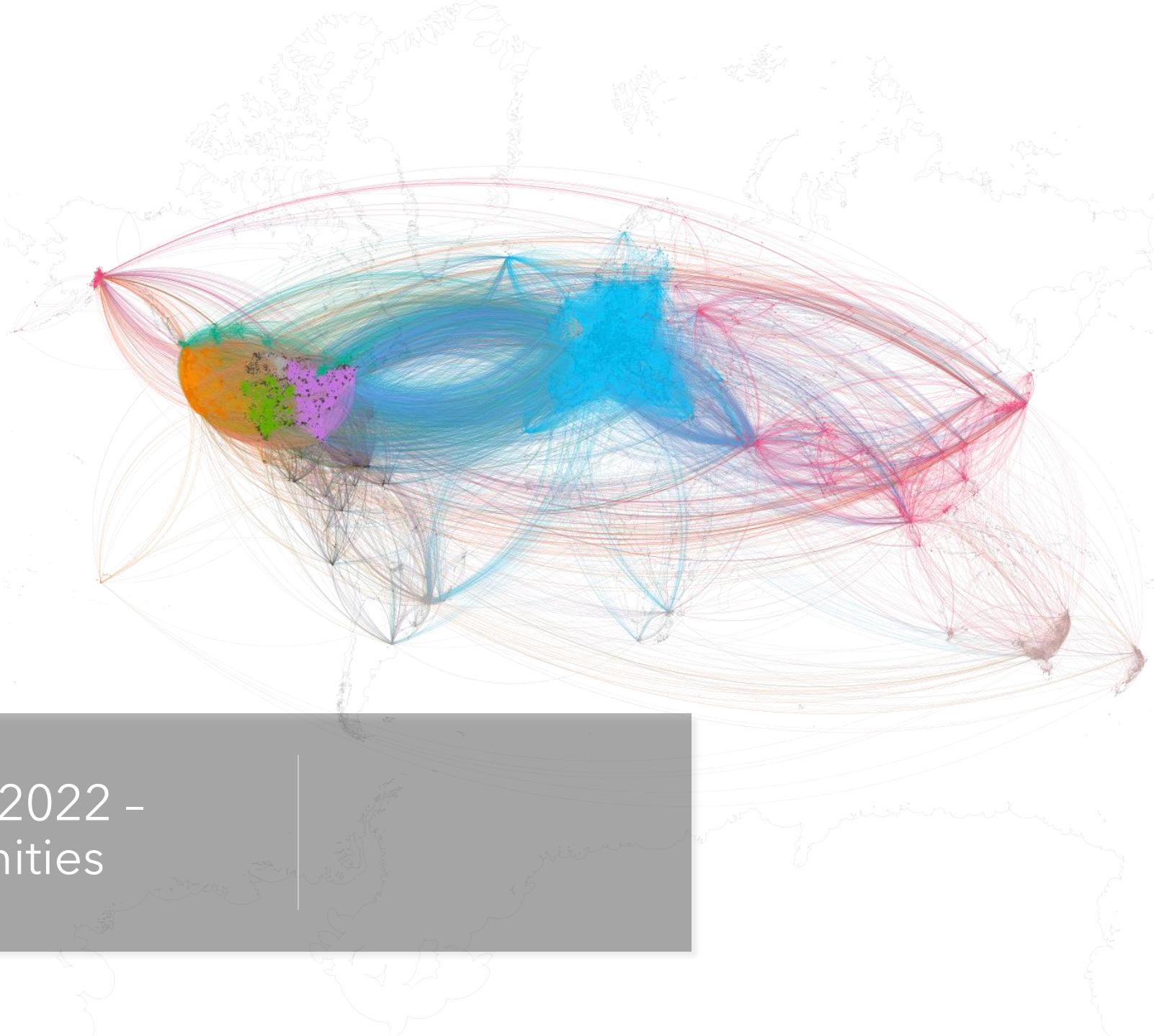
Airports 2020 - Communities





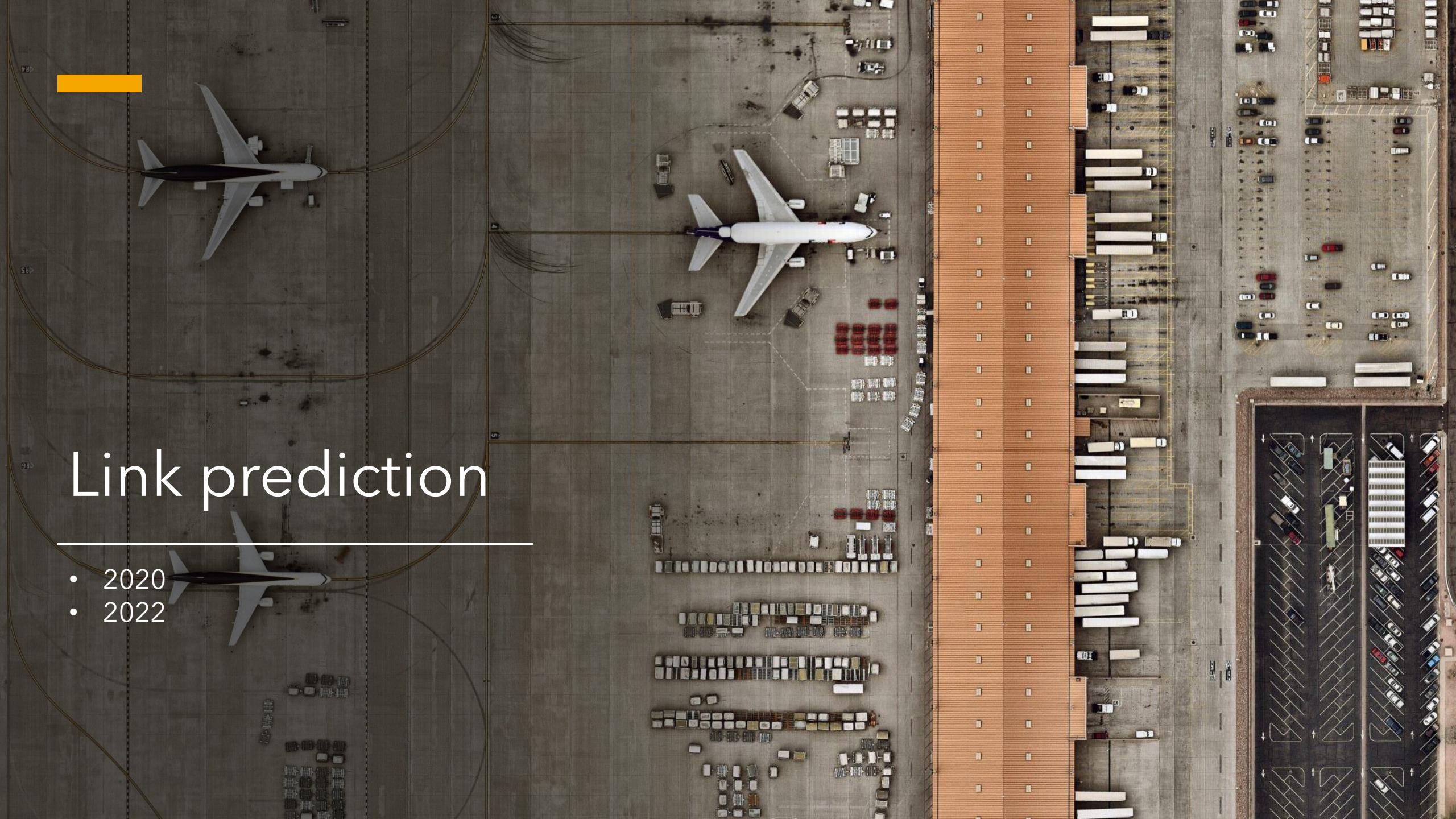
Airports 2021 - Communities

Airports 2022 - Communities



Link prediction

- 2020
- 2022



Link prediction - Description



Link prediction has been performed on the dataset of the years 2020 and 2022.



Following the same reasoning of the project, the weight of all edges is set to 1, so we do not consider multiple flights among two airports.



Multiple methods of link predictions have been used: resource allocation, preferential attachment and common neighbors.

Link prediction - 2020

Common
neighbors:

KABI	KONM
KIAH	KT85
KFRH	KONM
KIAH	KHLR
KAMT	KARR
KARR	KSOP
KABI	KPRC
KDNS	KONM
KONM	KTDF
KARR	KDKX

Preferential
attachment:

KABI	KONM
KIAH	KT85
KFRH	KONM
KIAH	KHLR
KAMT	KARR
KARR	KSOP
KABI	KPRC
KDNS	KONM
KONM	KTDF
KARR	KDKX

Resource
allocation:

KHON	KSPS
KSCK	KSPS
KIAH	KFCM
KIAH	KGMJ
KPAM	OMDW
CYML	CYYT
KCJR	SBKP
KPEX	KSDC
KHFD	KSBN
K4V9	KIAH

Link prediction - 2022

Common
neighbors:

27OK	KOMH
KARW	KDLS
KEUG	27OK
KDDH	KJEF
KDDH	KFNL
KDDH	KLKU
KDDH	KMQS
KSAT	KT74
KPHX	KT74
KDDH	KLVK

Preferential
attachment:

27OK	KOMH
KARW	KDLS
KEUG	27OK
KDDH	KJEF
KDDH	KFNL
KDDH	KLKU
KDDH	KMQS
KSAT	KT74
KPHX	KT74
KDDH	KLVK

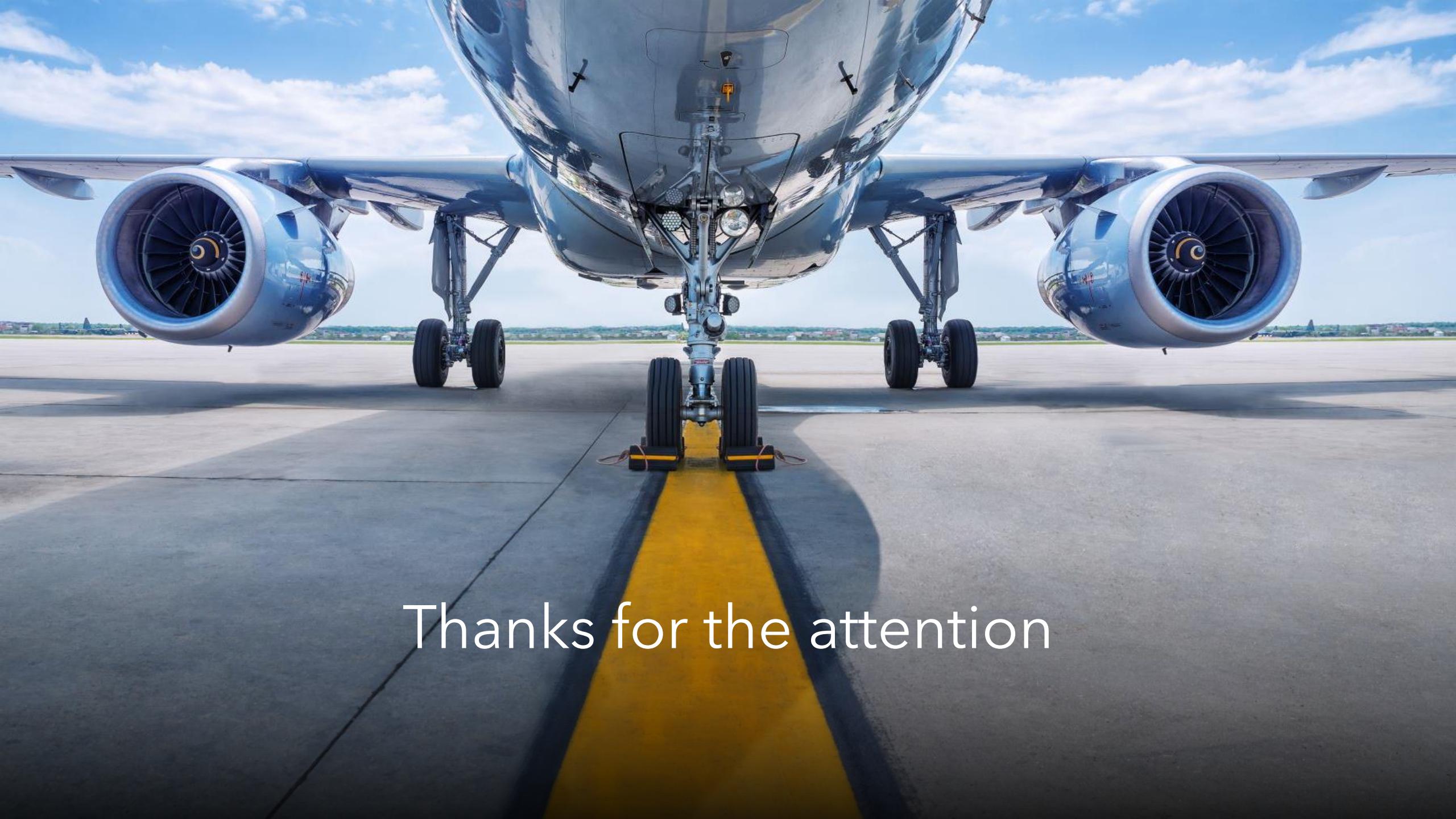
Resource
allocation:

27OK	KOMH
KDDH	KIPT
KDDH	KSWI
KDCA	KLHQ
KDDH	KLKU
KD98	KP13
KOMH	KSCD
KDDH	KLHZ
KDDH	KT17
K3MY	KSNL

Link Prediction

Prediction from 2020

Prediction from 2022



Thanks for the attention