



# Diaspora Mapping / Engagement Database List

NamSor, 2022

Personal names reflect cultural and linguistic identity. The geographic distribution of personal names of a specific origin in professional, scientific or other databases can be used as a proxy for conducting migration studies. Although this method does not pretend to provide 100% accuracy, [NamSor](#) can produce new and interesting data in a cost effective manner that complement other quantitative and qualitative methods. NamSor can also be instrumental in cost effective diaspora mapping / engagement strategies (for FDI / economic development, talent acquisition, real-estate, PPP / diaspora financing, startups & innovation ...)

## Contents

Diaspora Mapping / Engagement Database List.....	1
Introduction .....	2
Database List.....	3
Business Databases.....	3
Scientific Databases .....	4
General Databases .....	5
Residential Databases .....	5
Artists Databases.....	5
NamSor Licence Terms and DPA.....	5
Self-service on <a href="#">namsor.app</a> .....	5



## *Introduction*

[NamSor SAS](#) is a European vendor of specialized 'big data' mining software to conduct various types of analysis on personal names. Onomastics is a branch of sociolinguistics that can be applied to mine 'big data' and categorize personal names according to various taxonomies (such as gender, linguistic and cultural origin). We started software development in 2012 and since 2014, NamSor has proposed this technology of applied onomastics for diaspora mapping and targeted diaspora engagement.

In January 2019, NamSor launched a new version of the software, NamSor API V2 – which can be used a self service by researchers to process large volumes of data. Currently, NamSor has supported dozens of high-impact research projects from the Worlds' top universities.

NamSor also provides specific consulting services :

- Diaspora Mapping : data mining and processing of open databases or commercial database, in order to create geographic or logical maps and statistics about the Diaspora (by country/region or by sector/job level) after anonymization of the data ;
- Diaspora Engagement : processing and identification of potential Diaspora members in specific regions / business areas, in order to participate to certain programs and capacity building activities that benefit both the origin country and the host country (such as: attracting foreign direct investment, boosting entrepreneurship, building-up science and education via international collaboration, ...) on behalf of a national or international organization;

This document lists the main databases NamSor has experience in working with.



## Database List

Each database contains personal names, as well as additional information. Technically, any database can be processed using NamSor API and create a new database controlled by the Client.

- Open Databases can generally be processed for Diaspora Mapping and Statistics to produce a document/report.
- Open Databases can generally be processed for Diaspora Engagement to produce a new database. The resulting database contains names and personal information. As such, it should be controlled by the Client and handled appropriately with respect to data privacy, communication with prospects (opt-in, opt-out).
- Commercial Databases generally require a specific License between the Commercial Vendor and the Client, so NamSor can do the processing on behalf of the Client (using the Client's API Key or web access).
- Contact enrichment databases are used to prepare an outreach program. After manually reviewing the select profiles for Diaspora engagements (ex. influencers, prominent diaspora members ...), these databases are used to find contact information and prepare the CRM database.

## Business Databases

Business databases are generally used for B2B marketing, corporate sales activities, risk management, financial and legal matters.

Database Name	Scope	Description	Homepage	License
ZoomInfo	Global	150 million B2B profiles, globally with a specific focus on Northern America and Western Europe	<a href="https://www.zoominfo.com/business/our-data">https://www.zoominfo.com/business/our-data</a>	Commercial*
UK Companies House Register	UK	The list of UK Company Directors, Secretaries or People with significant control (PSC, ie. Beneficial owners)	<a href="https://beta.companieshouse.gov.uk/">https://beta.companieshouse.gov.uk/</a>	Open Data
BVD ORBIS	Global	Orbis has information on more than 365 million companies across the globe.	<a href="https://www.bvdinfo.com/en-gb/our-products/data/international/orbis">https://www.bvdinfo.com/en-gb/our-products/data/international/orbis</a>	Commercial
BVD DIANE	FR	Diane contains comprehensive information on companies in France.	<a href="https://www.bvdinfo.com/en-gb/our-products/data/national/diane">https://www.bvdinfo.com/en-gb/our-products/data/national/diane</a>	Commercial
CrunchBase	Global	Database of Start-ups	<a href="https://data.crunchbase.com/docs/getting-started">https://data.crunchbase.com/docs/getting-started</a>	Commercial
Fullcontact	Global	Global contacts database	<a href="https://www.fullcontact.com/">https://www.fullcontact.com/</a>	Commercial
People Data Labs Inc.	Global	Global contacts database (3.1B+ people records)	<a href="https://www.peopledatalabs.com/">https://www.peopledatalabs.com/</a>	Commercial

\*For ZoomInfo, NamSor can provide Diaspora Mapping and Statistics for a fixed fee, but the Client needs to enter a Commercial License with ZoomInfo and obtain an API Key for accessing the detailed profiles (incl. contact information). The cost per profile is in range 0.2\$-5\$ depending on volumes.

## Scientific Databases

Scientific publication databases and patents databases are generally used for scientific collaboration and advancement of science and technology.

Database Name	Scope	Description	Homepage	License
ORCID	Global	the global scholar register, with about ~6 million scientists	<a href="https://orcid.org/about">https://orcid.org/about</a>	Open Data
PUBMED/PMC	Global	the largest scientific open database focused on Life Sciences	<a href="https://www.ncbi.nlm.nih.gov/pmc/">https://www.ncbi.nlm.nih.gov/pmc/</a>	Open Data
DOAJ	Global	DOAJ (Directory of Open Access Journals) is a community-curated online directory with about ~7 million scientists.	<a href="https://doaj.org/">https://doaj.org/</a>	Open Data
CORD-19	Global	COVID-19 Open Research Dataset (CORD-19) w/ about ~10,000 Coronavirus scientists	<a href="https://pages.semanticscholar.org/coronavirus-research">https://pages.semanticscholar.org/coronavirus-research</a>	Open Data
NCIS	Global	Novel Coronavirus Information Center	<a href="https://www.elsevier.com/connect/coronavirus-information-center">https://www.elsevier.com/connect/coronavirus-information-center</a>	Open Data
DBLP	Global	At the end of March 2020 dblp provides bibliographies for almost 2.5 million computer scientists.	<a href="https://dblp.uni-trier.de/">https://dblp.uni-trier.de/</a>	Open Data
Elsevier SCOPUS	Global	Elsevier's scientific publications database with ~16 million author profiles.	<a href="https://www.elsevier.com/solutions/scopus">https://www.elsevier.com/solutions/scopus</a>	Commercial
Clarivate Web of Science (WOS)	Global	Clarivate flagship database of scientific publications with wide coverage.	<a href="https://www.webofknowledge.com">https://www.webofknowledge.com</a>	Commercial



## General Databases

General information databases include mass media and social media monitoring, provide insights about public personalities of any importance (VIPs, influencers).

Database Name	Scope	Description	Homepage	License
Factiva	Global	General VIPs cited in Newspapers globally.	<a href="https://professional.dowjones.com/factiva/">https://professional.dowjones.com/factiva/</a>	Commercial
webhose	Global	Media monitoring API with VIPs cited globally.	<a href="https://webhose.io/">https://webhose.io/</a>	Commercial
Twitter	Global	Twitter historical database	<a href="https://developer.twitter.com/en/docs/tutorials/choosing-historical-api">https://developer.twitter.com/en/docs/tutorials/choosing-historical-api</a>	Commercial

## Residential Databases

Residential databases can be used to estimate the demographics of the population, when no other geographic or statistical information is available.

Database Name	Scope	Description	Homepage	License
FR Residential	FR	French landlines directory (Pages Blanches)	<a href="https://www.pagesblanches.fr/pagesblanches">https://www.pagesblanches.fr/pagesblanches</a>	Open Data / Commercial
US Voters Lists	US	US Voters Lists (Over 190 million U.S. voters available) via NationBuilder	<a href="https://nationbuilder.com/voterfile">https://nationbuilder.com/voterfile</a>	Open Data / Commercial

## Artists Databases

Artists databases provide useful information about countries soft powers in cultural industries (such as the film industry, or music).

Database Name	Scope	Description	Homepage	License
IMDB	Global	IMDb is the world's most popular and authoritative source for movie, TV and celebrity content	<a href="https://www.imdb.com/">https://www.imdb.com/</a>	Open Data / Commercial
Discogs	Global	the largest online music database	<a href="https://www.discogs.com/">https://www.discogs.com/</a>	Open Data

## NamSor Licence Terms and DPA

[https://v2.namsor.com/NamSorAPIv2/assets/pdf/201803\\_NamSor\\_API\\_Terms\\_v007.pdf](https://v2.namsor.com/NamSorAPIv2/assets/pdf/201803_NamSor_API_Terms_v007.pdf)

[https://v2.namsor.com/NamSorAPIv2/assets/pdf/NamSor\\_Data\\_Processing\\_Agreement\\_v007.pdf](https://v2.namsor.com/NamSorAPIv2/assets/pdf/NamSor_Data_Processing_Agreement_v007.pdf)

## Self-service on namsor.app

We've made it possible for our users to self-service their diaspora mapping and engagement program, by licensing these databases and filtering data through NamSor app (<https://namsor.app/>).