

Identifying KOLs for Bright Life's NeuroShield


TEAM

Abhishek Krovvidi

krovvidi@purdue.edu

linkedin.com/in/abhishek-krovvidi/

Harshraj Jadeja

hjadeja@purdue.edu

linkedin.com/in/harshraj-jadeja/

Shivam Mishra

mishr211@purdue.edu

linkedin.com/in/shivam-mishra-b485b4142/

PURDUE UNIVERSITY | Mitchell E. Daniels, Jr. School of Business

ProcDNA

AGENDA

01

About the Product

NeuroShield: Cutting-Edge Protection for Brain Health, leveraging advanced neuroscience to shield against degenerative diseases and enhance cognitive function

02

Methodology

Encompasses data collection, preprocessing and EDA, culminating in the development of a scoring model for KOL selection

03

Data

Crafted by web-scraping with API's, focusing on influential and customer-centric attributes for KOLs

04

Exploratory Data Analysis

Conducted to identify trends, patterns, and correlations in data for feature selection and optimization

PURDUE UNIVERSITY | Mitchell E. Daniels, Jr. School of Business

ProcDNA

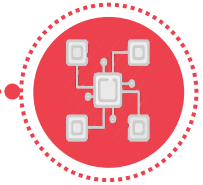
AGENDA

05

Data Preprocessing

Involves cleansing, structuring and refining the dataset to enhance the quality and reliability of the data

06



Model Building

Entails constructing a robust, scoring model and a network model to depict KOL influence

07

Final Recommendation

Implemented and managed the model through a dashboard, ensuring ongoing accuracy and adaptability to market changes

08

Future Scope

Address current limitations and expand the model's capabilities for broader, future applications

PURDUE UNIVERSITY | Mitchell E. Daniels, Jr. School of Business

ProcDNA

01

Project Scope

Defining project outline

Business Problem

Identify influential Key Opinion Leaders (KOLs) in the US Neurology market to support Bright Life company's new product launch strategy.

Project Outcome

A strategic plan identifying and engaging top KOLs in Neurology, enhancing product endorsement and adoption within the medical community.

Analytics Problem

Develop a methodology to analyze datasets, including clinical trials, research publications, leadership roles, and financial engagements, to identify and rank KOLs.

Key Deliverables

A presentation detailing the analysis methodology, KOL profiles, and a comprehensive list of identified KOLs, supported by analysis files/codes.

ProcDNA

01

Why?

Introducing the Product

"Parkinson's disease is the second most common neurodegenerative disease and is an important societal issue and global priority."

See Series page 283

There are 1 million living with Parkinson's in the USA

60,000 Americans are diagnosed with Parkinson's disease each year, and this number does not reflect the thousands of cases that go undetected

The combined direct and indirect cost of Parkinson's, including treatment, social security payments and lost income from inability to work, is estimated to be nearly $25 billion per year

Medication costs for an individual person with PD average $2,500 a year, and therapeutic surgery can cost up to $100,000 dollars per patient

ProcDNA

01

About the Product

Introducing the Product

NeuroShield: Early Parkinson's Intervention Therapy

Real-time Monitoring and Insights Generation

Early Detection Algorithm

Localized TMS Therapy

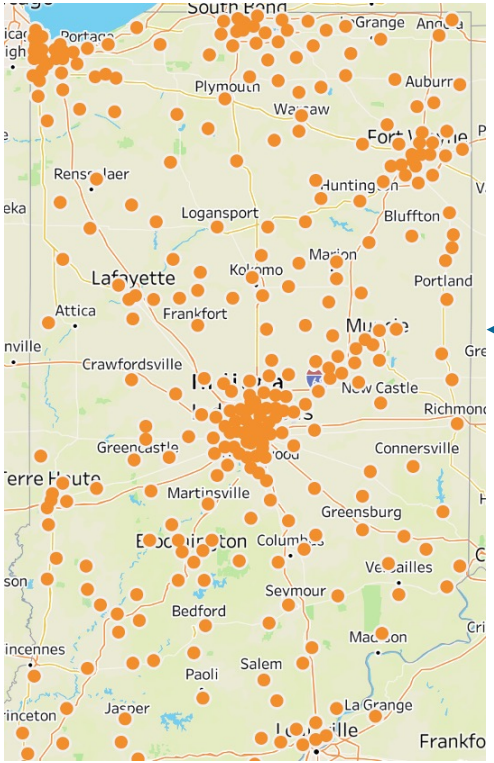
Wearable and Non-Invasive

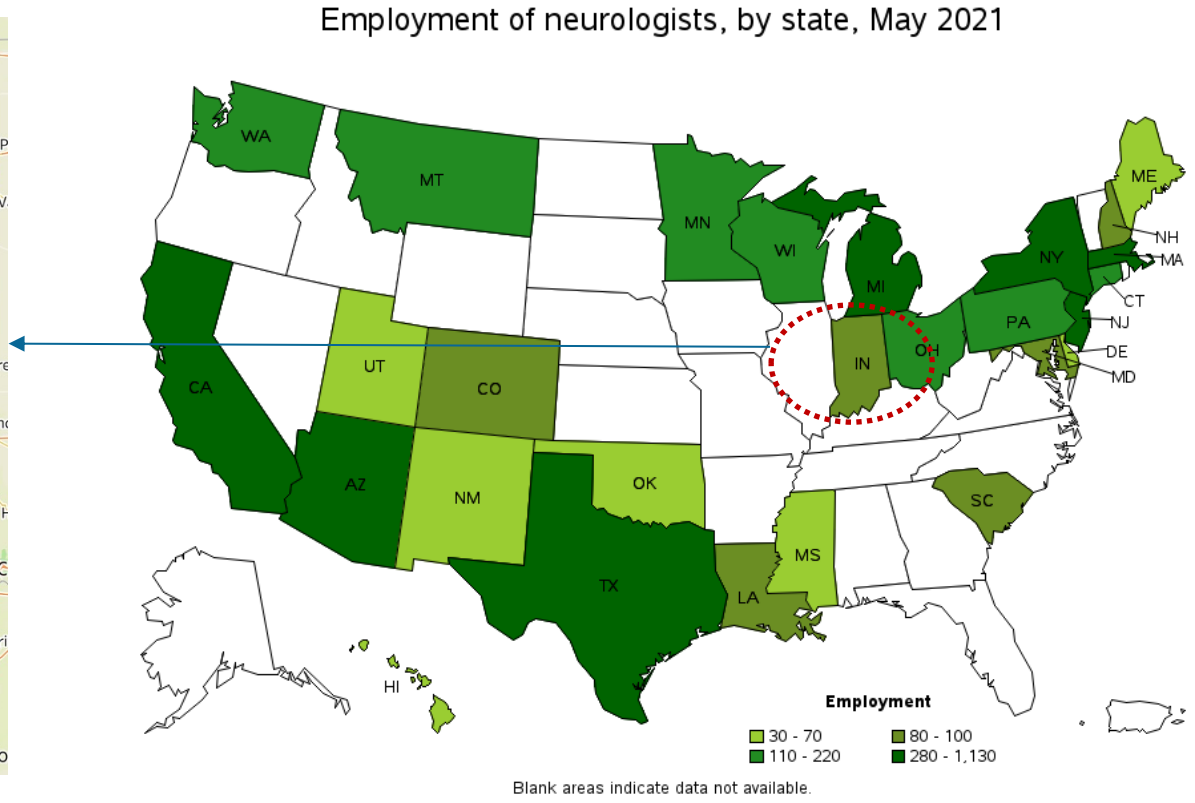
ProcDNA

01

Target Audience

Introducing the Product



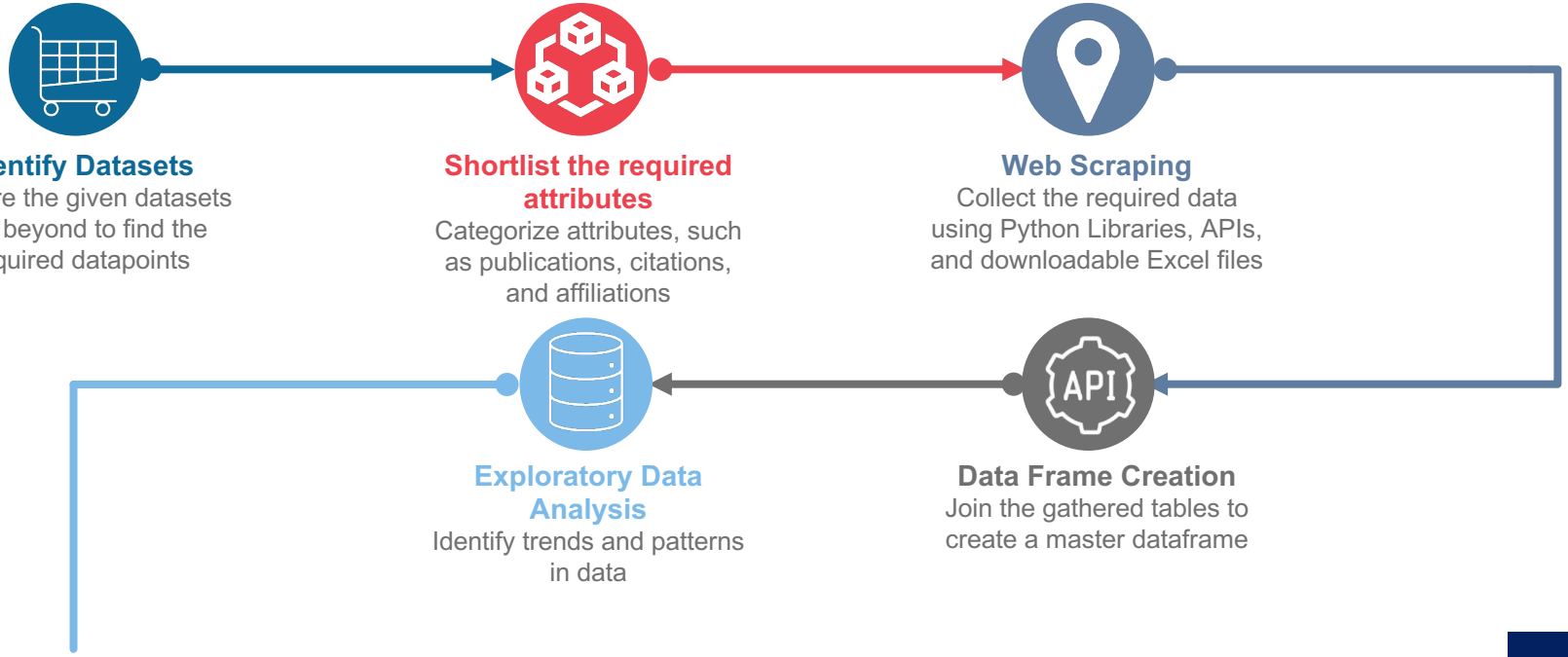


ProcDNA

02

Methodology

Analytical process framework



Identify Datasets
Explore the given datasets and beyond to find the required datapoints

Shortlist the required attributes
Categorize attributes, such as publications, citations, and affiliations

Web Scraping
Collect the required data using Python Libraries, APIs, and downloadable Excel files

Data Frame Creation
Join the gathered tables to create a master dataframe

Exploratory Data Analysis
Identify trends and patterns in data

ProcDNA

Data Preprocessing
Filter data by removing / imputing the records to get a cleaner dataset

Feature Engineering
Altering the required features as it suits best to create a model

Modeling
Rank all the profiles based on a scoring metric and select top 10 profiles

Recommendations
Shortlist the accumulated profiles further based on social reach and accomplishments to get the top 2 profiles

Future Scope
Explore how to scale this model to different fields of medicine

ProcDNA

03

Datasets

Collection of clinical trials data and payments data for a physician

Knowledge/Prior Experience

Research Journals
Pub Med

Clinical Trials
Clinical Trials . Gov

Research Payments
Open Payments Data

Affiliations
Web of Science

Social Influence

Twitter/LinkedIn
No of followers

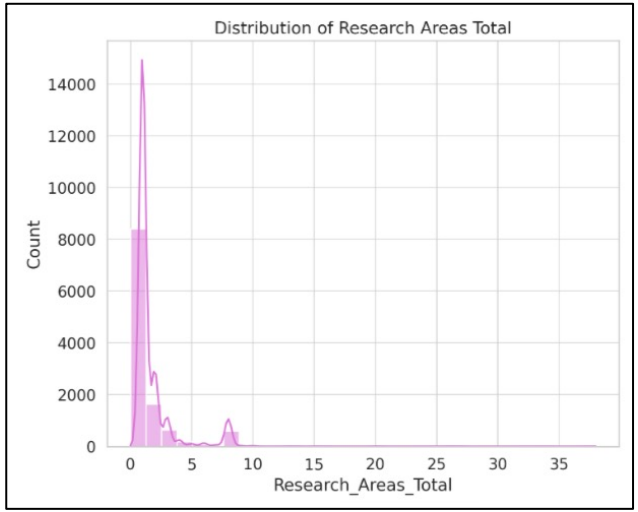
US News Ratings
Sentiments

ProcDNA

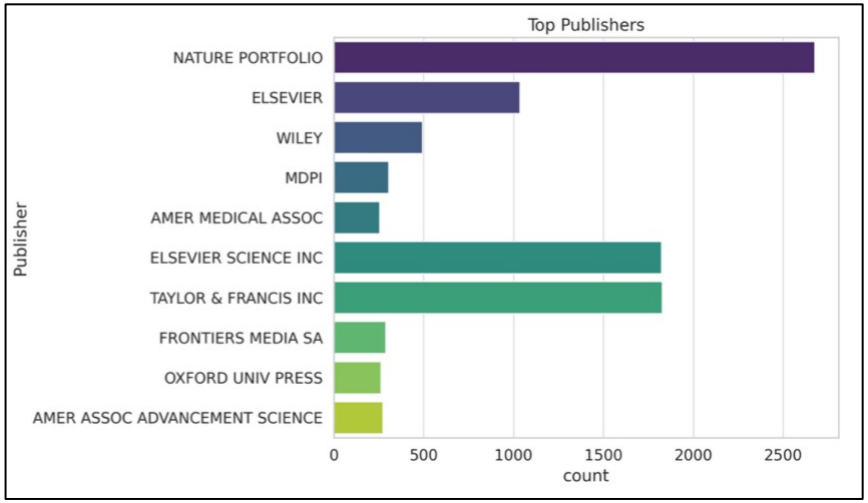
04

Exploratory Data Analysis

To identify trends, patterns, and correlations in data for feature selection and optimization



Visualizes the total number of research areas covered by the authors, indicating that most authors focus on a limited number of research areas, with a few exploring a broader range



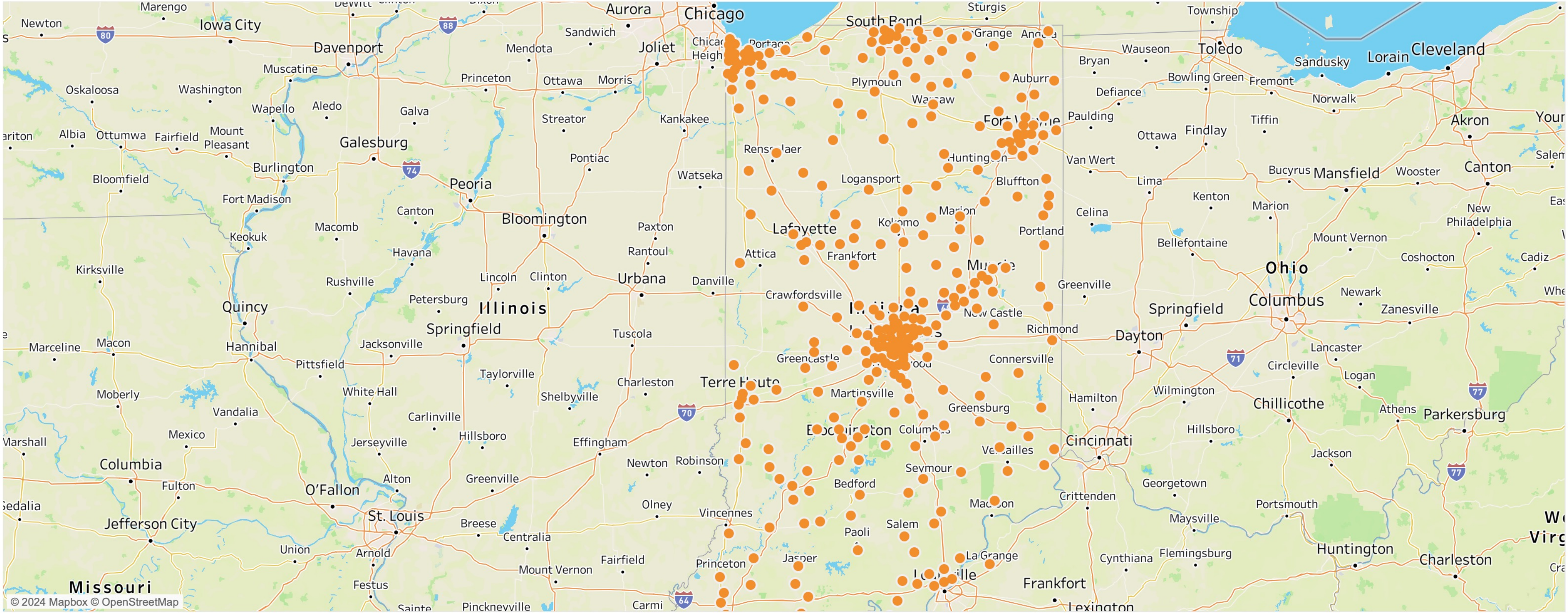
Provides a count of publications by the top publishers in the dataset, illustrating which publishers are most prevalent among your dataset's publications

PrecDNA

Exploratory Data Analysis for NeuroShield

Total Funding	Total Affiliations	Total Citations	Total Research Areas
398,664	2,319,516	3,427,219	20,225

Map of Available Doctors in Indiana State



05

Data Preprocessing

Enhances the quality and reliability of the data, paving the way for more accurate and meaningful insights

Why?

Enhance Data Quality

Removing irrelevant and highly correlated columns improves the overall quality and reliability of the dataset

Feature Engineering

By creating new features through effective bucketing and data manipulation, we ensured removing noise

Ensure Data Integrity

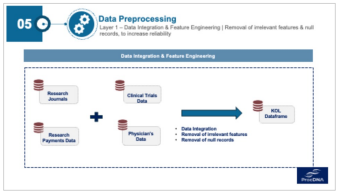
Removing nulls from columns helps in maintaining the integrity of the dataset, ensuring that the analysis is not skewed

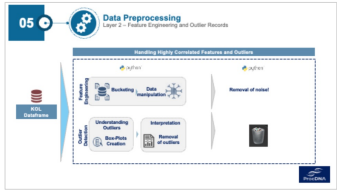
Removal of Outliers

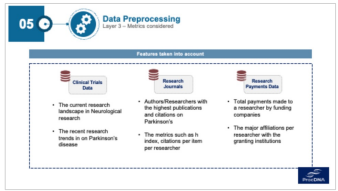
Removing outliers can enhance model accuracy by reducing noise and ensuring the training process focuses on the more representative patterns of the data

Layered Approach

Data Preprocessing Framework







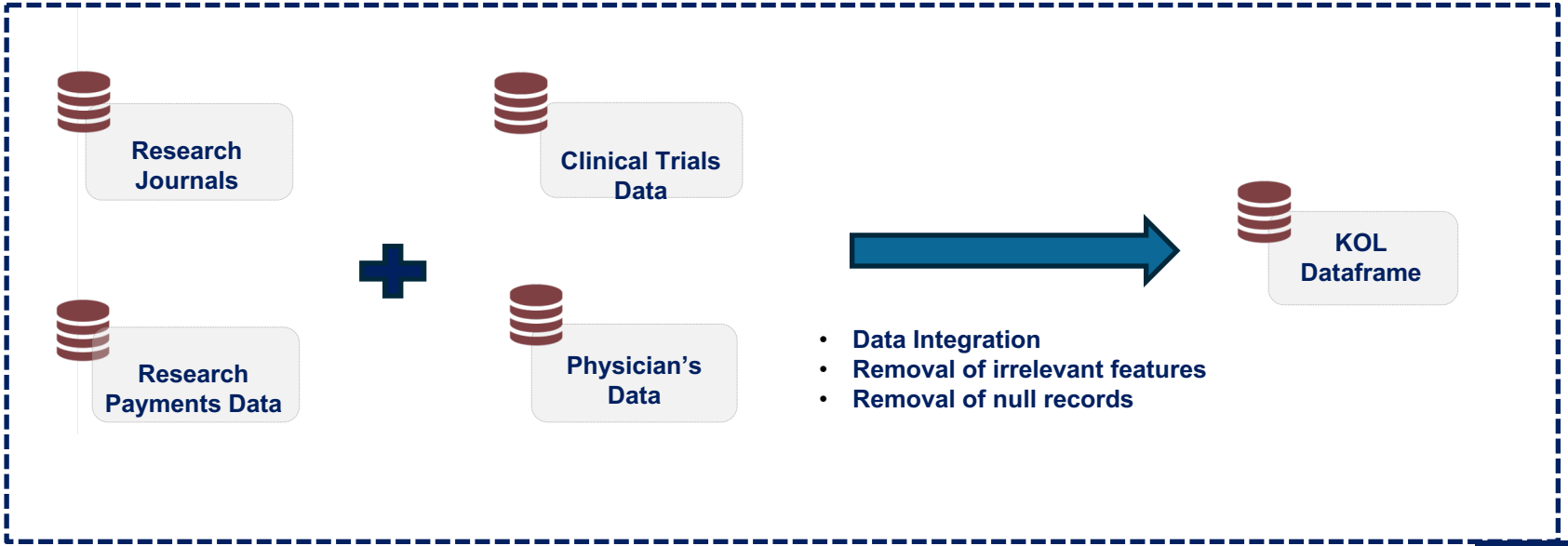
ProcDNA

05

Data Preprocessing

Layer 1 – Data Integration & Feature Engineering | Removal of irrelevant features & null records, to increase reliability

Data Integration & Feature Engineering



ProcDNA

05

Data Preprocessing

Layer 2 – Feature Engineering and Outlier Records

Handling Highly Correlated Features and Outliers

KOL Dataframe

python

Feature Engineering

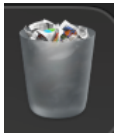
Bucketing → Data manipulation

Outlier Detection

Understanding Outliers → Box-Plots Creation → Interpretation → Removal of outliers

python

Removal of noise!



ProcDNA

05

Data Preprocessing

Layer 3 – Metrics considered

Features taken into account

Clinical Trials Data

- The current research landscape in Neurological research
- The recent research trends in on Parkinson's disease

Research Journals

- Authors/Researchers with the highest publications and citations on Parkinson's
- The metrics such as h index, citations per item per researcher

Research Payments Data

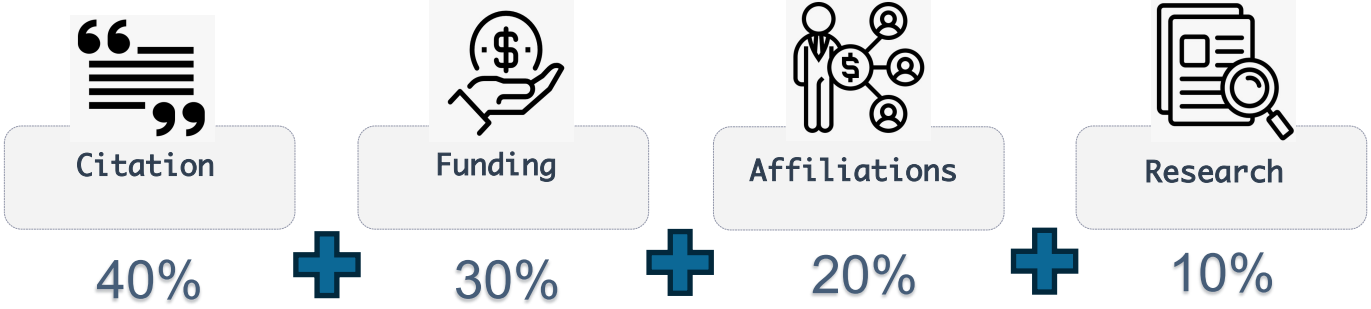
- Total payments made to a researcher by funding companies
- The major affiliations per researcher with the granting institutions

ProcDNA

06

Model Building

Stage 1 | Heuristic Scoring Model



- Total citations per KOL highlight their significance and ongoing impact within the neurology field
- Total funding and number of grants per KOL reflect the extensive financial support and trust from funding bodies
- Total affiliations per KOL showcase their extensive network and esteemed positions in the medical and academic communities
- The number of unique research areas per KOL demonstrates their broad interests and significant contributions to neurology

PrecDNA

07

Final Recommendation

Heuristic Score Ranking

Leadership Positions

Social Influence

KOL Mapping Influence Network

KOL Profiles

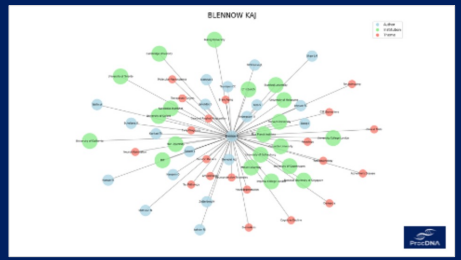
ProcDNA

BLENNOW KAJ

Professor/Chief Physician | University of Gothenburg
kaj.blennow@clinchem.gu.se

Specialty: Clinical Neurology

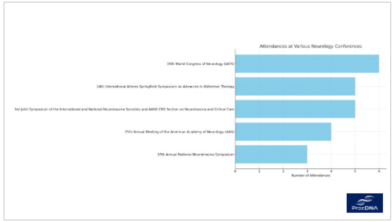
77.3 Influence Score



Key Research Areas

- Neurosciences Neurology
- Psychiatry
- Cell Biology
- Geriatrics Gerontology

Leadership



Interests

- Good health & well-being
- Gender Equality
- Zero Hunger

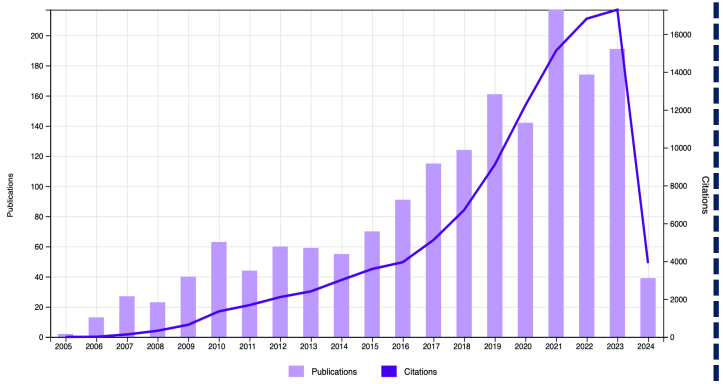
Affiliations

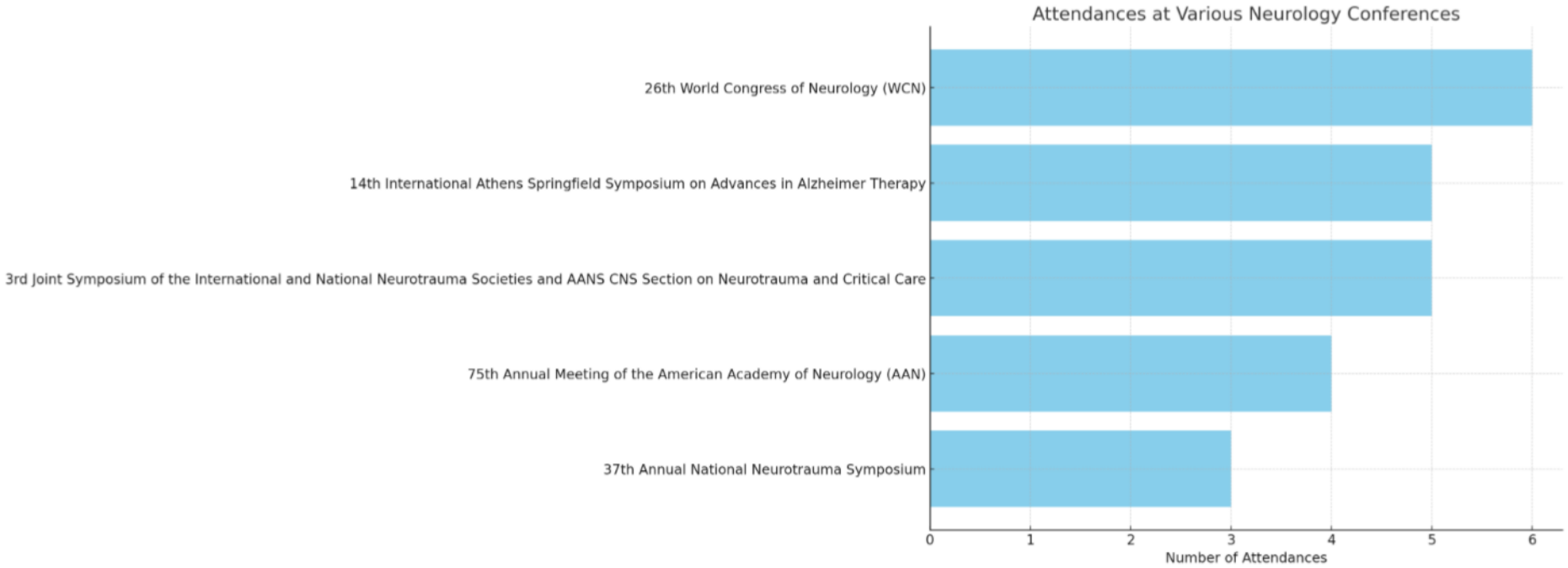
Research Journals

1,710 Publications

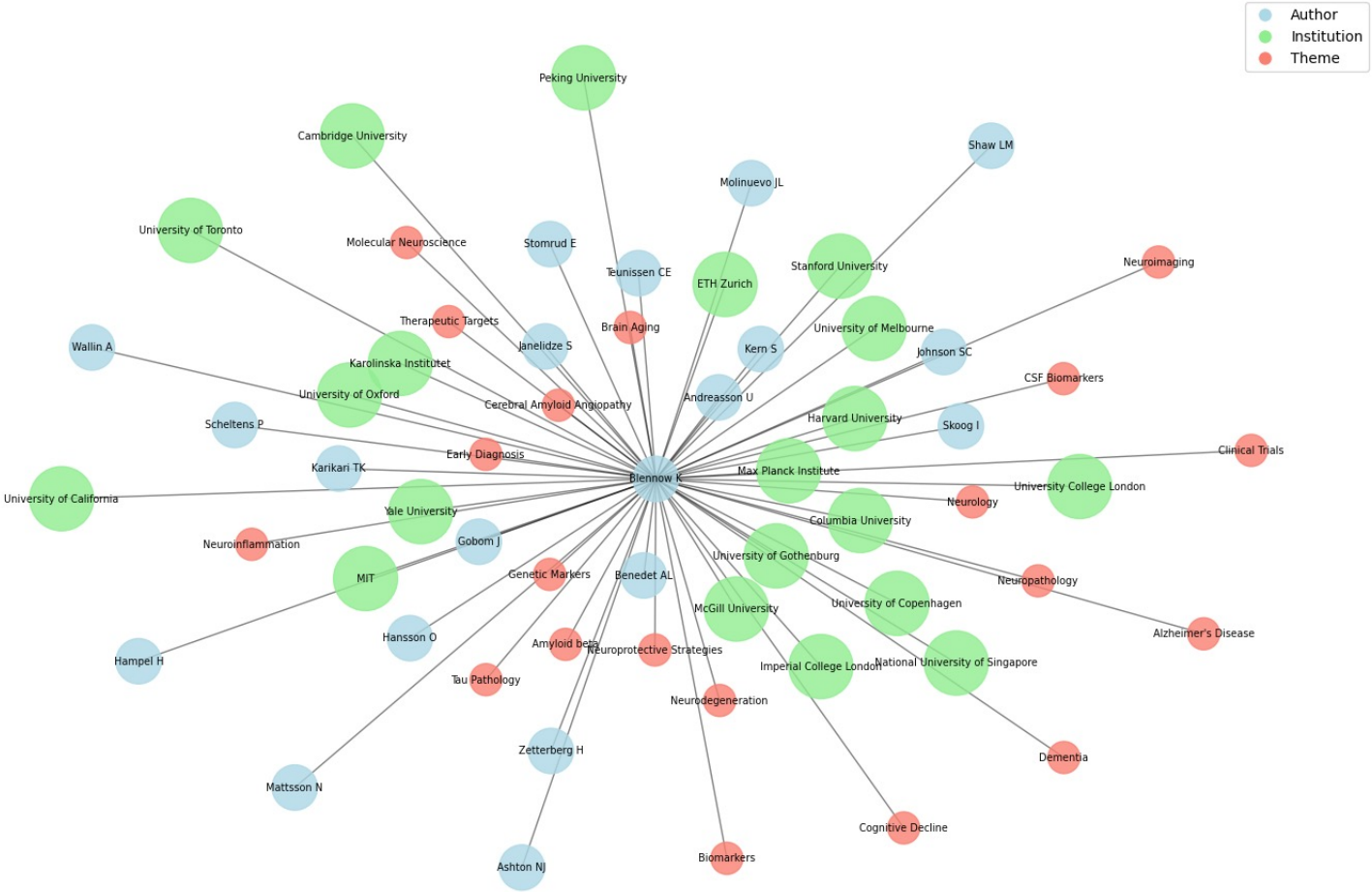
105K Citations

144 H-Index




Attendances at Various Neurology Conferences
26th World Congress of Neurology (WCN)
14th International Athens Springfield Symposium on Advances in Alzheimer Therapy
3rd Joint Symposium of the International and National Neurotrauma Societies and AANS CNS Section on Neurotrauma and Critical Care
75th Annual Meeting of the American Academy of Neurology (AAN)
37th Annual National Neurotrauma Symposium
0
1
2
3
4
5
6
Number of Attendances

BLENNOW KAJ


Author
Institution
Theme
Peking University
Cambridge University
Shaw LM
Molinuevo JL
University of Toronto
Molecular Neuroscience
Stomrud E
Teunissen CE
ETH Zurich
Stanford University
Neuroimaging
Therapeutic Targets
Brain Aging
University of Melbourne
Wallin A
Janelidze S
Kern S
Johnson SC
Karolinska Institutet
CSF Biomarkers
University of Oxford
Cerebral Amyloid Angiopathy
Andreasson U
Harvard University
Scheltens P
Skoog I
Early Diagnosis
Clinical Trials
Karikari TK
Max Planck Institute
University of California
University College London
Yale University
Neurology
Columbia University
Neuroinflammation
Gobom J
University of Gothenburg
MIT
Genetic Markers
Benedet AL
Neuropathology
McGill University
University of Copenhagen
Hansson O
Alzheimer's Disease
Amyloid beta
Neuroprotective Strategies
Hampel H
Imperial College London
National University of Singapore
Tau Pathology
Neurodegeneration
Dementia
Zetterberg H
Mattsson N
Cognitive Decline
Biomarkers
Ashton NJ

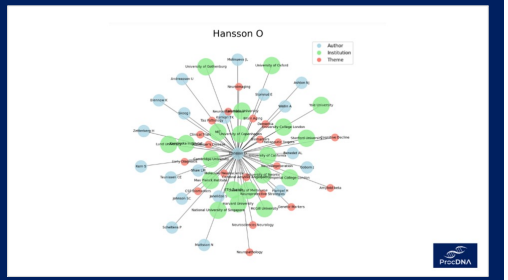
HANSSON OSKAR

2.4k+

Professor of Neurology| Lund University
https://twitter.com/OskarHansson9

Specialty: Neurology

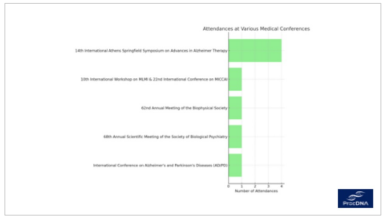
100 Influence Score



Key Research Areas

- Neurosciences Neurology
- Radiology
- Cell Biology
- Science Technology

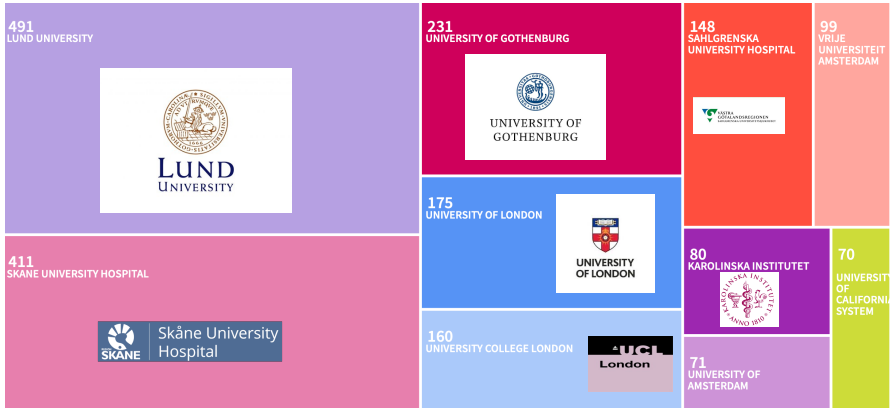
Leadership



Interests

- Good health & well-being
- Gender Equality

Affiliations

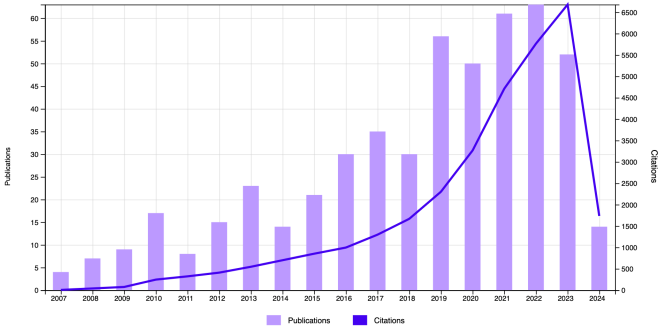


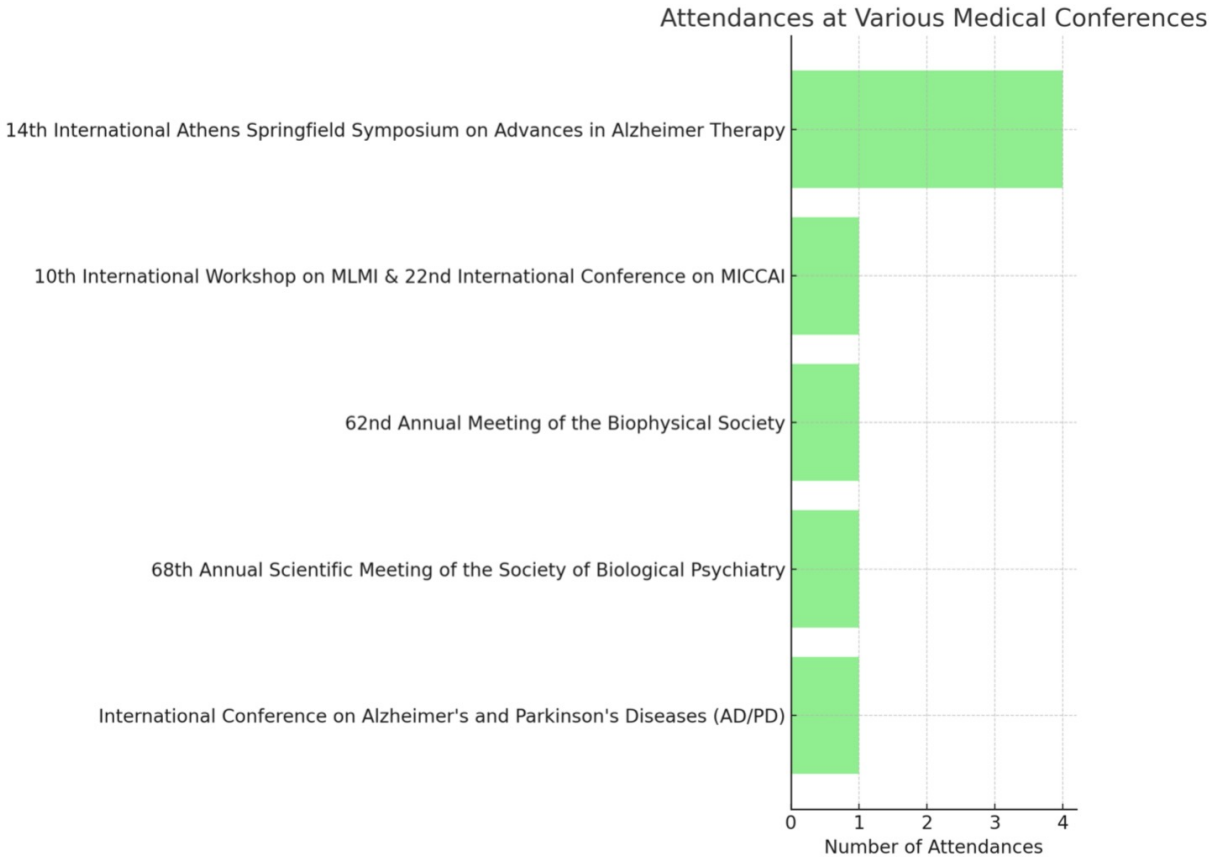
Research Journals

509 Publications

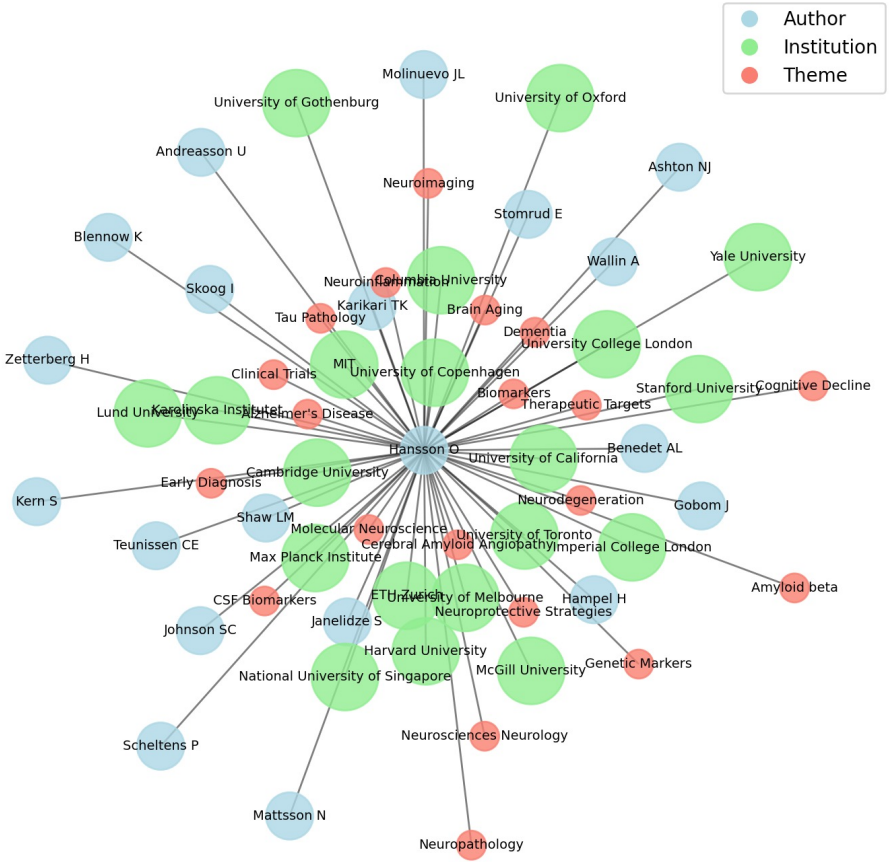
31K Citations

98 H-Index




Attendances at Various Medical Conferences
14th International Athens Springfield Symposium on Advances in Alzheimer Therapy
10th International Workshop on MLMI & 22nd International Conference on MICCAI
62nd Annual Meeting of the Biophysical Society
68th Annual Scientific Meeting of the Society of Biological Psychiatry
International Conference on Alzheimer's and Parkinson's Diseases (AD/PD)
0
1
2
3
4
Number of Attendances

Hansson O


Author
Institution
Theme
Molinuevo JL
University of Gothenburg
University of Oxford
Andreasson U
Neuroimaging
Ashton NJ
Stomrud E
Blennow K
Wallin A
Yale University
Columbia University
Skoog I
Karikari TK
Brain Aging
Tau Pathology
Dementia
University College London
Zetterberg H
MIT
University of Copenhagen
Clinical Trials
Biomarkers
Stanford University
Cognitive Decline
Therapeutic Targets
Lund University
Karolinska Institutet
Alzheimer's Disease
Hansson O
Benedet AL
University of California
Cambridge University
Early Diagnosis
Kern S
Neurodegeneration
Gobom J
Shaw LM
Molecular Neuroscience
Teunissen CE
University of Toronto
Cerebral Amyloid Angiopathy
Imperial College London
Max Planck Institute
Amyloid beta
CSF Biomarkers
ETH Zurich
University of Melbourne
Hampel H
Neuroprotective Strategies
Johnson SC
Janelidze S
Harvard University
Genetic Markers
McGill University
National University of Singapore
Scheltens P
Neurosciences Neurology
Mattsson N
Neuropathology

07

Validation/Accuracy Assessment

Timeliness

- Data arrival at a specified time interval
- Check for "late data" arrival

Uniqueness

- Check duplication of records at specified level of data (Primary key)

Accuracy & Validity

- Trend based anomaly detection for specified fields – flag anomalies outside statistical tolerance bands
- Categorical columns have a valid entry from a defined list of values
- Have business sense checks in place

Consistency

- Compare data between tables to ensure accurate numbers (use the aggregated view as benchmark)
- Comparison with Business forecasts

Completeness

- Null/missing value checks
- Outlier detection for the volume of data arriving on a day – compare the count of rows within a window of the last 30 days and send an alert if it's outside a defined limit

- **Timeliness, Uniqueness, Accuracy & Validity would be a part of general checks (that could be leveraged for all the datasets)**
- **Consistency & Completeness checks must be tailored/customized as per the specific datasets**

ProcDNA

08

Future Scope

Address current limitations and expand the model's capabilities for broader, future applications

01 Data Source Enhancement

Incorporate premium datasets for deeper insights and superior KOL profiling accuracy, moving beyond freely available internet resources.

02 Scope Expansion to Other Medical Fields

Extend the KOL identification framework to additional medical specialties beyond neurology, broadening the study's applicability and impact

03 Leveraging Advanced Visualization and ML Models

Enhance visibility and insights for KOL identification with advanced visualization techniques and leverage historical data in ML models for improved precision in data-driven approaches

ProcDNA

Thank You