Kleefstra syndrome data collection and the Rare Diseases Observatory

Tanja Zdolšek Draksler, PhD
tanja.zdolsek@ijs.si
HOW IS THE KS COMMUNITY COLLECTING DATA?

• Kleefstra Syndrome worldwide map: www.kleefstraworldmap.org

• GENIDA, collecting caregiver-reported data: https://genida.unistra.fr/

• RARE-X, collecting caregiver reported data - https://kleefstra.rare-x.org/

• AllStripes, collecting de-identified eHRs (only USA, Canada, UK) - IDefine USA
LESSONS LEARNED

• Raising awareness about the importance of data collection is crucial and still needed
• Need for learning and training for health literacy and digital literacy
• Data collection programs need to adapt to different users - mobile versions crucial, personalized for different users, better UX/UI, enabling different languages, data visualizations...
RARE 2030 RECOMMENDATIONS (FORESIGHT)

• No. 7: OPTIMISING DATA FOR PATIENT AND SOCIETAL BENEFIT

• Data used to its maximum to improve the health and well-being of people living with a rare disease.

• Lead by Eurordis

www.rare2030.eu/recommendations/
Enabling easy monitoring of rare diseases, providing information and knowledge to empower all: parents, clinicians, researchers, policy makers...

By combining different global data sources

- For monitoring global media news (2014-present): Event Registry system (news intelligence platform),
- For social media monitoring (2008-present): Twitter data (Twitter API for Academic Research)
- For scientific publications monitoring (2001-present): Medline, Microsoft Academic Graph and OpenAlex.
RARE DISEASES OBSERVATORY - Methodology

- Data extracted based on 20 specific keywords (16 rare diseases and 4 general terms)
- **Pilot version RDO** presents 16 neurodevelopmental disorders: *Kleefstra syndrome, Kabuki syndrome, Koolen-de Vries syndrome, Phelan-McDermid syndrome, Pitt Hopkins syndrome, Cornelia de Lange syndrome, Angelman syndrome, Fragile X syndrome, Dravet syndrome, Rett syndrome, SATB2 syndrome, Prader-Willi syndrome, SYNGAP1 syndrome, CTNNB1 syndrome, Wiedemann-Steiner syndrome, FOXG1 syndrome.*
- **ELK Stack** (Elastic Search, visualizations: Kibana)
- **Sentiment analysis:** open-source library NLTK - vader_lexicon for Twitter data and performed from EventRegistry for media news
Future plans

- Increase the number of monitored rare diseases (focus on NDDs stays)
- Custom made visualizations
- Real-time data feed
- New data sources?
- Improved UI/UX
- Other based on focus group and testing users feedback
16 MONITORED NDDs AT A GLANCE

Shows NDDs and their occurrence in global media news, social media, and scientific publications. The area of the circle represents the number of news, social posts, and scientific publications.
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors Name</th>
<th>Source Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonix Pharma Expedites Fibromyalgia and Chronic Migraine Programs</td>
<td></td>
<td>Investing.com South...</td>
</tr>
<tr>
<td>Tonix Pharma (TNXP) Expedites Fibromyalgia and Chronic Migraine Programs</td>
<td></td>
<td>StreetInsider.com</td>
</tr>
<tr>
<td>Tonix Pharmaceuticals Expedites Fibromyalgia and Chronic Migraine Programs</td>
<td></td>
<td>Green Stock News</td>
</tr>
<tr>
<td>Met Police officers guilty of gross misconduct over racist WhatsApp messages</td>
<td></td>
<td>The Independent</td>
</tr>
<tr>
<td>Texas A&amp;M Experts Achieve Groundbreaking Phase Angelman Syndrome Development</td>
<td></td>
<td>India Education.E...</td>
</tr>
<tr>
<td>Colchester police officer to run marathon for friend’s child fighting rare illness</td>
<td></td>
<td>The Gazette</td>
</tr>
<tr>
<td>Trente-sept équipes de futsal au 4e challenge Gard-Ardèche Manon Valla</td>
<td></td>
<td>midilibre.fr</td>
</tr>
</tbody>
</table>
MONITORING SCIENTIFIC PUBLICATIONS
There are over 7000 rare diseases, and the majority are under researched. Estimated 400 million people around the world have a rare disease, but only 5% of rare diseases have a chance of treatment. That is why rare diseases need to be researched.

The Rare Diseases Observatory enables easy monitoring of rare diseases. By combining different global data sources, we provide the information an knowledge needed to empower all: patients, researchers and clinicians. Our initial focus is on rare neurodevelopmental disorders (NDDs).

WE ARE MONITORING RARE DISEASES
THROUGH DIFFERENT DATA SOURCES & INDICATORS

Feedback via GoogleForm
https://forms.gle/aiTt8FaR8