

# Может ли "голос пациента" в Интернет-среде стать источником данных реальной клинической практики?

Раннее выявление и исследование пути пациентов с тяжелыми заболеваниями с использованием технологий понимания естественного языка

Ирина Ефименко, Semantic Hub (основатель, global CEO),  
к.ф.н. в области компьютерной лингвистики, доцент

© Randy Glasbergen  
glasbergen.com



**"You have to learn about thousands of diseases, but  
I only have to focus on fixing what's wrong with ME!  
Now which one of us do you think is the expert?"**

Роль  
пациента:  
не только  
«объект  
заботы»

# Форматы вовлечения пациентов в принятие решений: международные практики

- **NICE** (Великобритания): **Программа вовлечения граждан** в разработку инструкций, стандартов и клинических рекомендаций - Public Involvement Program (создаются комиссии и комитеты, куда пациенты или лица, осуществляющие уход, входят **наравне с профессионалами**: они могут поделиться своим опытом и **обратить внимание на детали, неочевидные для профессионалов**. На сайте размещаются **вакансии**)
- **ЕМА** (Европа): Привлечение пациентов в качестве членов Правления и членов научных комитетов. Участие в т.ч. в изменении дизайна КИ
- **FDA** (США): **Сессии слушания пациентов** - Patient Listening Sessions: (регулярные закрытые для публики телеконференции, проводимые в неформальной обстановке, в состав которых входят пациенты и/или их близкие и сотрудники FDA – могут собираться в т.ч. **по инициативе пациентского сообщества**)
- **FDA** (США): Различные форматы **рабочих групп с участием пациентов (включая временное трудоустройство – в рамках которого предоставляется доступ к конфиденциальной информации)**  
– в т.ч. по вопросам создания новых моделей взаимодействия в системе здравоохранения, усовершенствования норм и **экспертизы новых медицинских продуктов**. Примеры: **Patient Engagement Collaborative, FDA Patient Representative Program, Patient-focused drug development, Patient Engagement Advisory Committee**

# «Голос пациента» в интернете

Обращаюсь к вам с очень срочным и сложным вопросом: **18 июня** этого года моя мама (**66 лет**) была прооперирована по поводу **РМЖ (первично-множественный синхронный рак). Радикальная мастэктомия по Маддену**. Диагноз после операционный: **первично-множественный синхронный рак. Рак правой молочной железы 1 А ст., pT1cN0M0 G1. Люминальный тип А. Рак левой молочной железы 1 А ст., pT1cN0M0 G1. Люминальный тип А. ИГХ которая проводилась до операции: справа Ki 10%, слева справа Ki 19%.** За месяц до и сразу после операции назначили **Анастрозол**, который она и пьет. К сожалению, **результаты нам готовили 1.5 месяца!** В процессе нам говорили, что анализы хорошие и, **скорее всего, ХТ не будет и мы 1.5 спокойно ждали, но когда приехали за заключением, нам сказали, что рекомендуют 4 курса химии, но решение за районным онкологом!** А время то уже прошло! Пока ждали запись к районному онкологу съездили в **2 клиники Медси на консультацию, где и заведующая химиотерапевтическим отделением, и консилиум врачей в клинике на Боткинской однозначно сказали ХТ не делать**, т.к. не видят смысла в ней. Считают, что **риски ХТ выше чем от нее возможный результат.** Когда попали к **районному онкологу она сказала, что химию делать надо и для нее показанием является выписка из больницы.** Я вас умоляю в срочном порядке, если это возможно, посмотреть выписку и высказать свое мнение, т.к. либо нам срочно начинать курс химии, либо нет... **ФРАГМЕНТ**



# От неструктурированной информации к структурированным данным (базе знаний)

- «Мне 60 лет, РМЖ T2N0M0. Сейчас пройдено 6 курсов ХТ: 4 курса доксорубицин + эндоксан и 2 курса паклитаксела. Решила перепроверить результаты ИГХ и теперь очень сожалею, что не сделала этого раньше. Отсутствие Er и Pr подтвердилось, но также обнаружили расхождения: Ki 67 – 80% (по первому заключению было 60%) и Her2 – 0 (по первому заключению 3+)...»



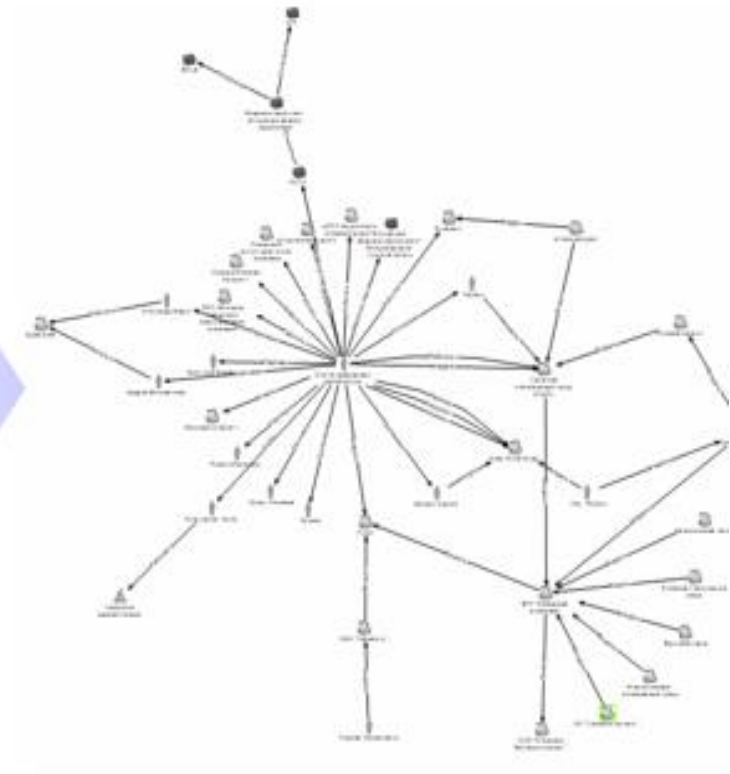
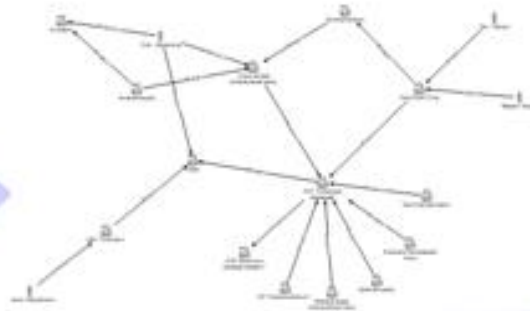
Semantic Hub 

Helping pharma see around the corner

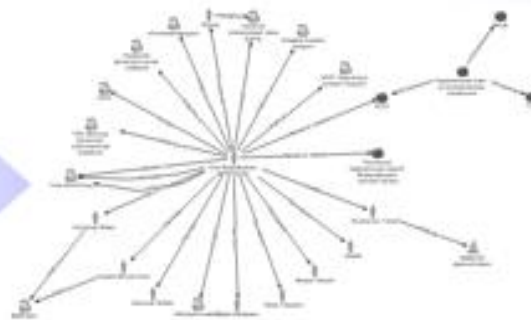
[https://www.flaticon.com/free-icon/dna\\_946375](https://www.flaticon.com/free-icon/dna_946375) (спасибо за иконки!)

# От неструктурированной информации к структурированным данным (базе знаний)

*My daughter is 2 years and 6 months old this year. She was born in full term. She is the second child (the first child is 8 years old, and she is OK). The daughter was able to stand when she was 8 months old and walked when she was 1 year and 2 months old. The walking posture is not normal (toes are on the ground, valgus foot), she is now 2 years and 6 months old, she is not able to stand freely, she needs support, thighs' muscles are thin and weak, this is mainly relevant for lower limbs (previously, the thighs were quite normal. When she was 1 year and 6 months, she had a low-grade fever for 1 week.*




*is to triple negative? If it is, no I'm oestrogen positive. I've been stage 4 for 3 years, got dx 2 years ago but had pain in my ribs for a year and docs kept telling me it was costocondritis. I'm on epirubicin ( I think lol ). I get it once a week for 2 weeks then a week off. I'm the same, we used to go all over but hardly go anywhere now. I'm so tired all the time and been really depressed lately. Think steroids haven't helped only had a 3 week course but they've knocked hell out of me. Gave me XXX cause I was getting pain from my liver and it was inflamed. However just after I felt pain which was something I wouldn't wish on anyone .it was all night long I couldn't even walk to the toilet...I hope you get chance to get out a bit*



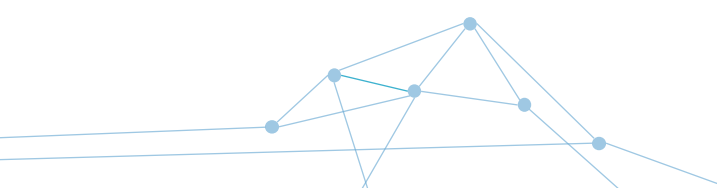
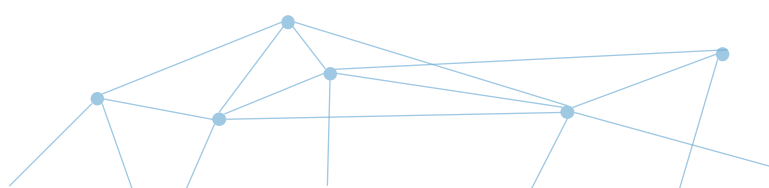
# От базы знаний – к инсайтам



\*Hospitals provided for each relevant city, town or region



# Данные, которые вы не найдете в другом месте *(даже при высоком качестве источников)*

- Качество жизни
  - Бремя заболевания и терапии (в некоторых случаях – единственный источник фармакоэкономических данных)
  - Критерии успешности терапии в глазах пациентов
  - Инсайты в области безопасности и эффективности (факты и восприятие)
  - Драйверы и барьеры, страхи пациентов
  - Проблемы доступа (которые никогда не появятся в официальных данных...)
  - Stakeholder mapping
- 
- 



# Технологии понимания ЕЯ

## Максимально неструктурированные данные

- «Я бросил принимать этот препарат, так как эффекта – НОЛЬ»
- «Зачем тратить деньги на эти дорогие таблетки? Все равно как мертвому припарка...»
- «Я сказала врачу: «Не помогает мне ваш Х... Только желудок болит». Мне тут недавно посоветовали масло шалфея, так что, девочки, лучше буду мазаться им»



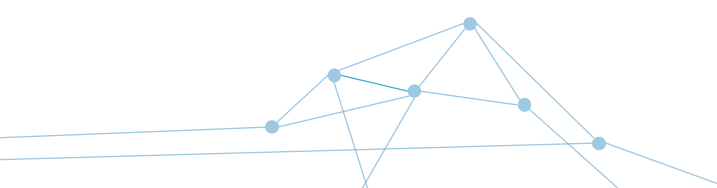
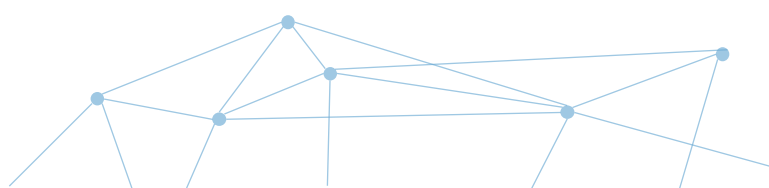
*Отсутствие приверженности по причине отсутствия эффективности*





# Технологии понимания ЕЯ

## Максимально неструктурированные данные

- ЭТИ ТИПЫ СТАЛИ ЕСТЬ НА СКЛАДЕ
  - ПАЛИТСЯ ЛИ ХОБЛ ФЛЮШКОЙ?
  - ПЕДЯ РЕКОМЕНДУЕТ ИЗБЕГАТЬ КАБАКОВ
  - 10 ЛЕТ НА АНАЛОГЕ, ТЕПЕРЬ - ПАРЕНИЕ
- 
- 

# Портфель показаний

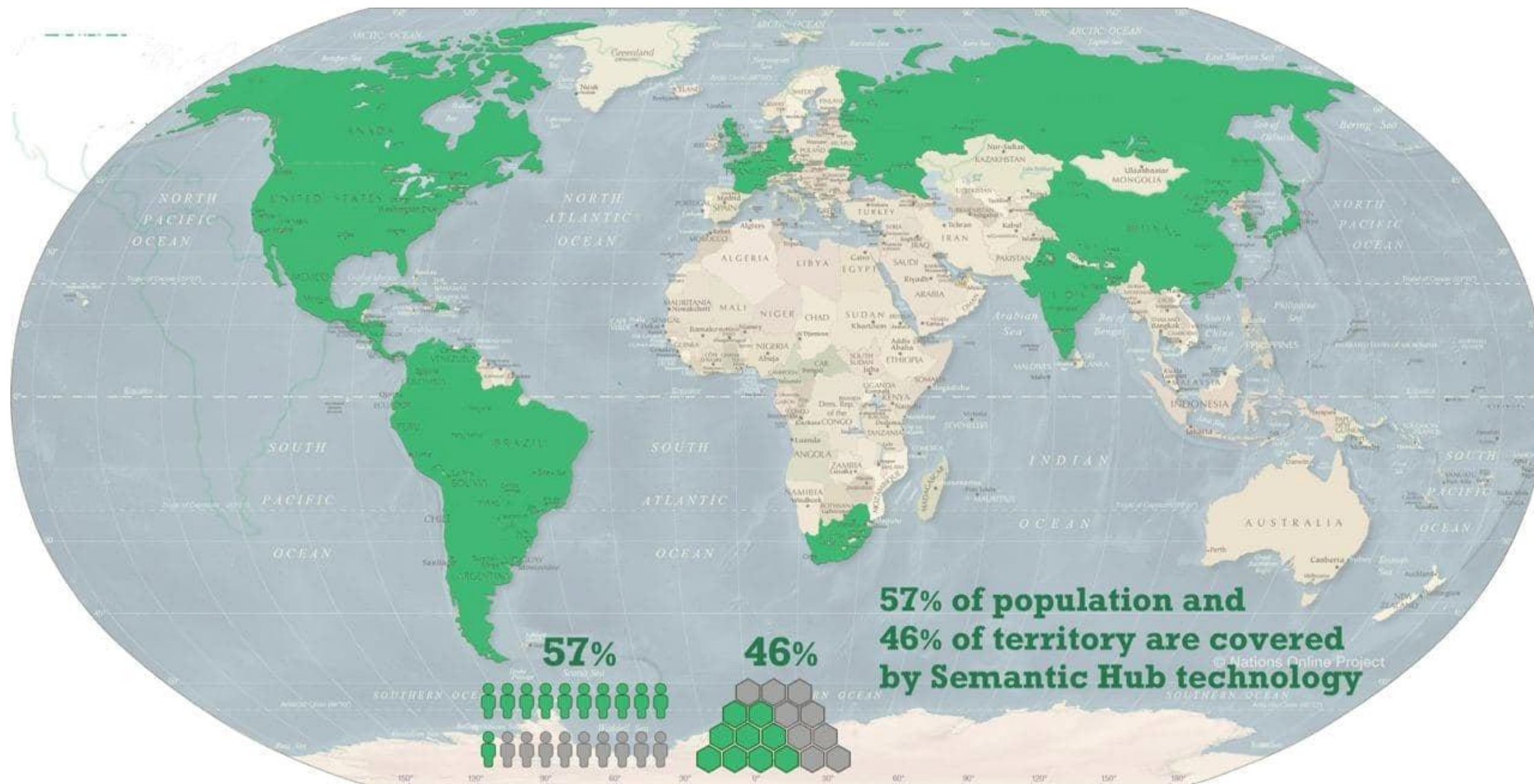
60+ редких и других тяжелых заболеваний:

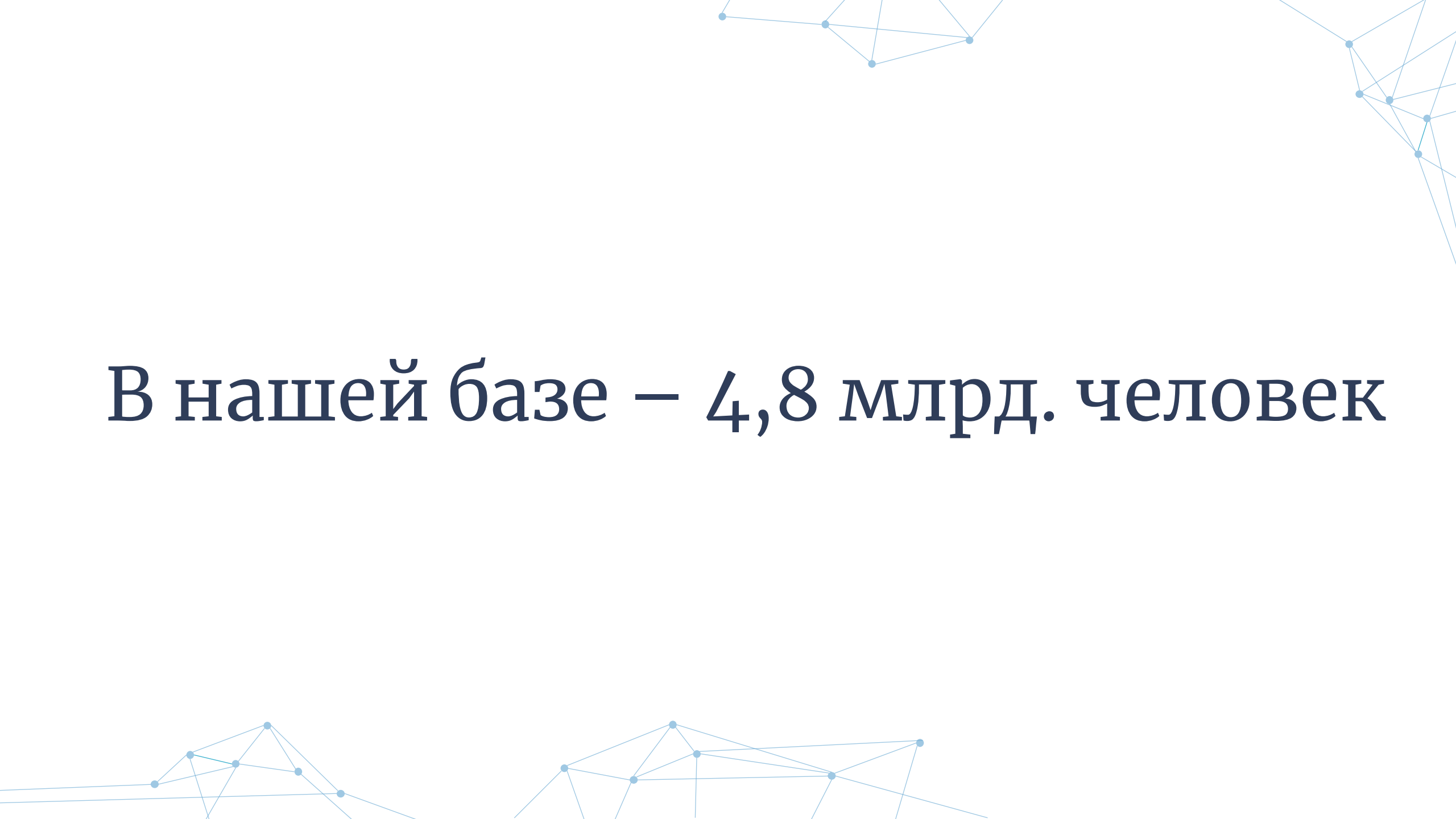
- **Онкология:** ~20
- **Редкие заболевания:** ~30
- **Другие тяжелые заболевания, выборки, соответствующие сложной системе критериев:** 10+
- **Раннее выявление недиагностированных редких пациентов**
- **Анализ опыта / пути пациента**
- **Поиск пациентов для КИ**
- **Анонимизация по GDPR**
- **Перевод с «пациентского» языка на клинический**

## Терапевтические области

- Oncology incl. rare tumors
- Neurology
- Hematology
- Autoimmune diseases
- Immunology and allergy
- Dermatology
- Cardiology
- Respiratory diseases
- Endocrinology
- Lysosomal storage diseases
- Ophthalmology
- Gastroenterology and hepatology
- Nephrology
- Psychiatry
- Rare infectious diseases

# 60+ успешных проектов в 30+ странах



The image features abstract geometric patterns in the corners, consisting of thin blue lines connecting small blue dots to form various polygonal shapes. These patterns are located in the top-left, top-right, and bottom-left corners of the slide.

В нашей базе – 4,8 млрд. человек



# Голос пациента как источник RWD (Novartis)

## Euretina 2021 Virtual Abstracts

Type Free Paper

### Title

Semantic analysis of patient voice in wet AMD and diabetic retinal disease: Patients' journey and healthcare experience

### Purpose

The purpose of the study is to analyze Patient Voice (messages from patients with wet age-related macular degeneration (wAMD) and diabetic retinal disease (DRD), or their caregivers) from open Internet resources in the Russian Federation, using artificial intelligence (AI) methods: technologies for automated analysis of unstructured natural language texts, incl. semantic technologies. The study provides valuable information, not only on patients' characteristics and treatment patterns in real-world management of retinal diseases, but also on patients' attitude to the disease, diagnostic and treatment-related procedures, their needs and barriers to treatment. The need for such data is explained by the fact that there are no available comprehensive reviews of country-specific patients' 'journeys' assessed from patients' perspective and analyzed by extracting insights from 'big data'. Moreover, recent publications show that the healthcare system is ready to recognize such information as an important source of real-world data (RWD). The intensive development of the virtual environment boosted by COVID-19 pandemic also highlighted the responsibility of healthcare stakeholders to develop digital solutions for different groups of patients (apps, remote monitoring tools, web-

### Presenter

Olga Zaytseva Russian Federation

### Co-Author 1

Vladimir Neroev

### Co-Author 2

Yurii Zhulev

### Co-Author 3

Vera Petrakovskaya

### Co-Author 4

Anna Paleeva

### Co-Author 5

Maria Trifonova

### Co-Author 6

Irina Efimenko

### Co-Author 7

# Голос пациента как источник RWD (Servier)

## #51: A descriptive analysis of patient characteristics from Russian internet postings focused on CRC (iPatient study)

Khromitskaya Y.V.,<sup>1</sup> Paleeva A.,<sup>2</sup> Ivankov A.,<sup>2</sup> Shulepina E.V.,<sup>1</sup> Korovkina N.Y.,<sup>1</sup> Trifonova O.P.,<sup>1</sup> Saburkina N.,<sup>1</sup> Grishin A.G.,<sup>1</sup> and Krotkevich A.<sup>1</sup>

1 – Servier Russia, Moscow, Russian Federation; 2 – Semantic Hub, Moscow, Russian Federation

### BACKGROUND AND OBJECTIVES

As healthcare in general, including oncology, is gradually becoming more patient-centered,<sup>1</sup> patients' perspective on their disease, its diagnosis and treatments, grows in relevance and significance. Recent advances in colorectal cancer (CRC) treatment have led to prolonged patient survival,<sup>2</sup> measured in years even at metastatic stage. That, together with growing incidence in Russia,<sup>3</sup> has resulted in an increased exposure of non-healthcare professionals to the problems related to cancer care and rehabilitation. The objective of the present study was to clarify the characteristics, treatment pathways and unmet needs in Russian patients with CRC

### METHODS

iPatient is a big-data retrospective observational study based on automatic web-scanning of social media, and patients' and physicians' message boards across the entire Russian Internet. Semantic analysis using machine-learning approach and pattern-based rules of texts were carried out to identify and classify constructs related to CRC containing personal experience.

The whole process had 4 stages: 1) crawling and pre-processing (scanning Google and Yandex search engines and extracting potential data of interest); 2) data extraction from unstructured texts, using proprietary linguistic processor by Semantic Hub; 3) results generation; 4) data visualization

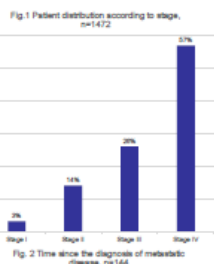
### RESULTS

Table 1

Overall population characteristics, n=2520	N(%)
Gender	
Male	759 (30%)
Female	1005 (40%)
Not specified	756 (30%)
Age bracket	
20-50 years	290 (12%)
51-70 years	550 (22%)
71 and older	216 (9%)
Not specified	1462 (59%)
Primary tumor localization	
Colon	1127 (44%)
Rectum	955 (38%)
Not specified	340 (13%)

Table 2

mCRC patient characteristics, n= 890	N(%)
Metastases localization	
Liver	605 (68)
Lungs	142 (16)
Peritoneum	99 (11)
Bones	80 (9)
Other	142 (16)
Number of metastatic sites	
1	401 (45%)
2	303 (34%)
3 and more	187 (21%)



In total, information on 2520 CRC patients was identified, coming from patients themselves in 56% of messages, and from their caregivers – in 44%. Geographically, Moscow was the most common region of origin (22% of patients), followed by Saint-Petersburg (10%) and Krasnodar region (4%). Patients demographic characteristics are represented in the Table 1 and show slight female predominance and relatively low number of patients older than 70. Of note, the vast majority of patients who specified their disease stage had Stage III-IV (83%, see Fig. 1), whereas according to statistics CRC in Russia is diagnosed at Stage I-II in half of patients.<sup>4</sup>

The characteristics of patients mentioning metastatic disease are summarized in the Table 2. The localization of metastases appears in line with known data, liver being most common, followed by lungs, peritoneum, and bones. 30% (n=267) of metastatic patients had been metastatic at the time of diagnosis.

Only minority of patients with metastases specified the duration of their disease: 146 out of 890 (16%), 72 (49%) of those having time since the diagnosis of metastatic disease more than 18 months (Fig.2)

Tumor mutation status (RAS, BRAF, MSI), ECOG performance status (PS) and defined concept of lines of therapy were almost never mentioned. Consequently, it was deemed impossible to properly analyze treatment sequences and therapy-related decision making.

### CONCLUSIONS

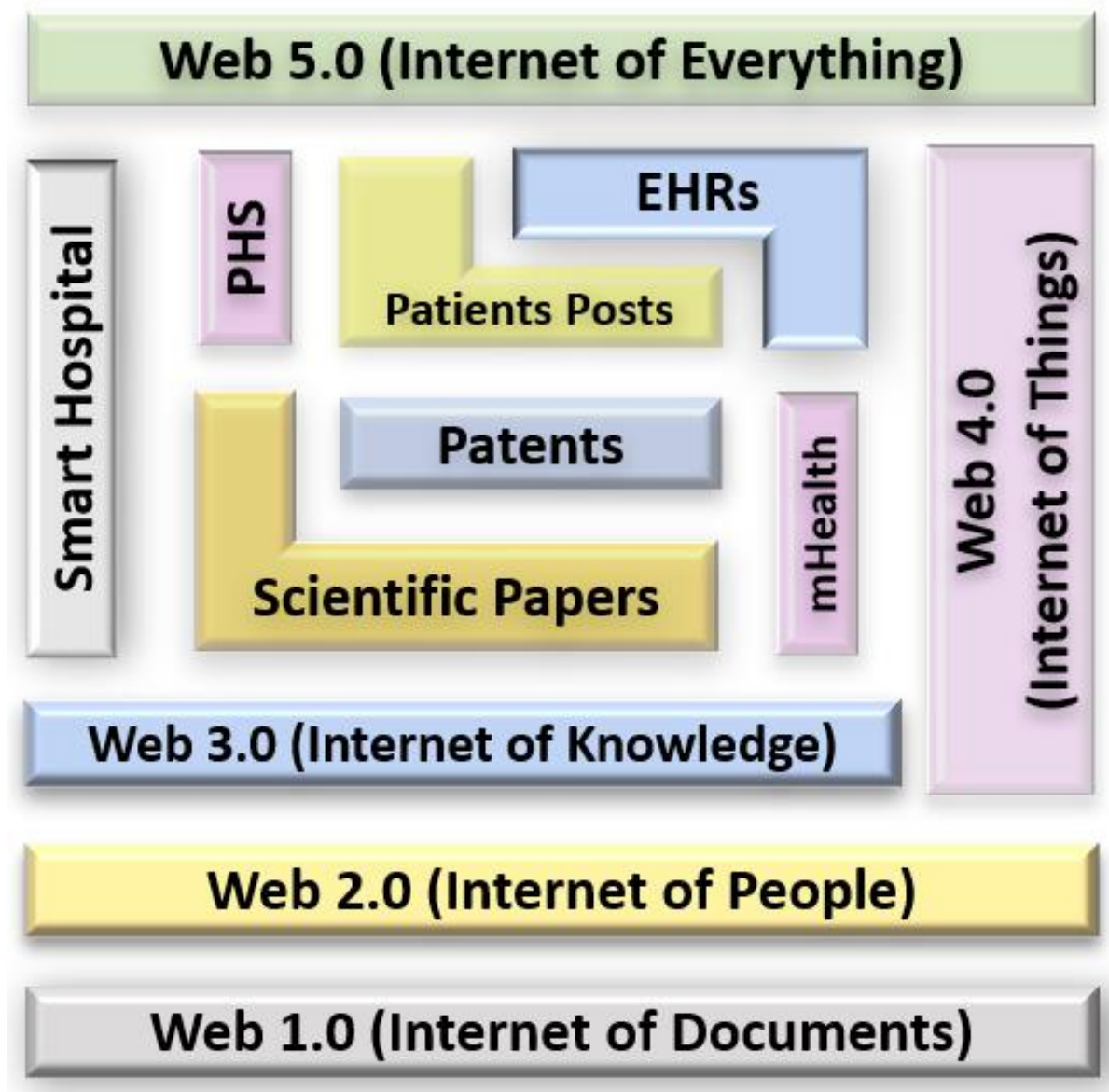
Analyzed CRC patient population is characterized by slight female predominance and relatively young age. One third of patients were diagnosed at metastatic stage. From the subset with mentioned number of metastatic sites, about half had low tumor burden (only 1 metastatic site). As for the duration of metastatic disease, almost half of those for whom it was specified had been metastatic for at least 18 months.

### References

1. Norberto J., Terk E., Sato B.D. et al. Overview on Patient Centricity in Cancer Care. *Front Pharmacol*. 2017; 8: 688.
2. Van Cutsem E., Cervantes A., Adam B., Sobrero A., Van Krieken J.H., Aderka D., Aranda Aguilar E., Bardelli A., Benson A., Bridle G., et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann. Oncol*. 2016; 27: 1154–1172.
3. Akhmetova M. Colorectal Cancer in Russia. *Eur J Health Econ*. 2010 Jan; 30 Suppl 1: S193-4.
4. Malignant Tumors in Russia in 2017 (Incidence and Mortality). Editors: A.D. Kaprin, V.Y. Stolyarsky, G.Y. Pavlov – P.A. Hereson. Moscow Oncology Research Center – Branch of FOMI MRCRC of the Ministry of Health of Russia, Moscow, 2018. [in Russian]

# ESMO-2020

# Большие данные: Web X.0



**НОВЫЕ ИСТОЧНИКИ  
RWD**



# We are looking forward to our cooperation!

*Semantic Hub*

**Please send us your request (an indication of your interest and the research questions)**

- so that we could prepare a detailed vision and the proposal for you,  
to help you better address the needs of your patients

**And please stay safe!**

*semantic-hub.com*  
*ie@semantic-hub.com*