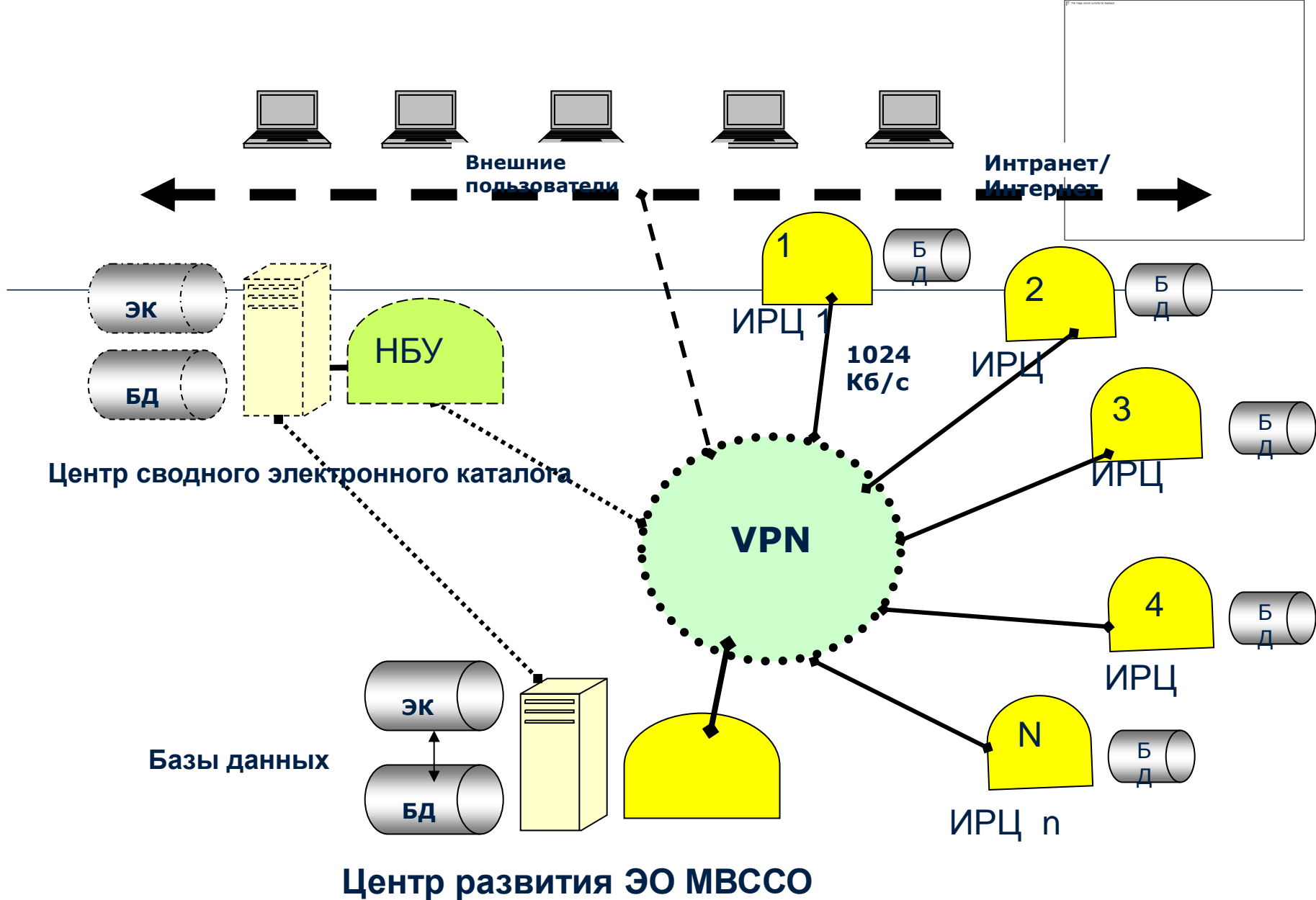


# **ОПТИМИЗАЦИЯ ПОИСКА В ЭЛЕКТРОННЫХ БИБЛИОТЕКАХ НА ОСНОВЕ ИНФЛЮЕНТНОГО АНАЛИЗА**

**Рахматуллаев Марат Алимович,  
профессор Ташкентского университета  
информационных технологий, д.т.н.**

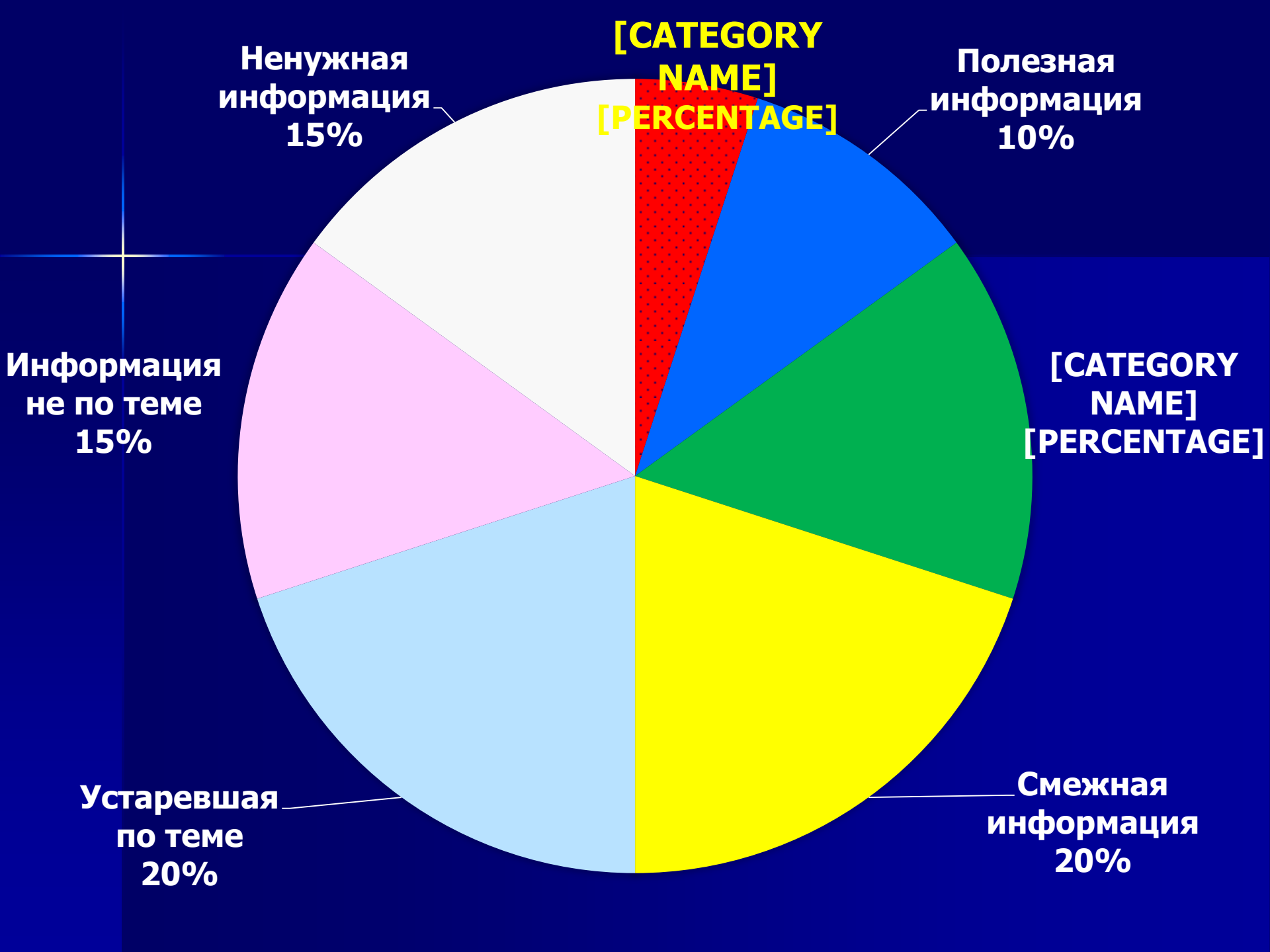




# Корпоративная сеть академических библиотек

## **Корпоративная сеть электронных библиотек характеризуется различными показателями.**

- **Оперативность поиска данных;**
- **Объем и содержание баз данных научной информации;**
- **Администрирование сети и баз данных;**
- **Организационная структура;**
- **Уровень квалификации кадров;**
- **Техническая инфраструктура**
- **Обеспечение информационной безопасности;**



**[CATEGORY NAME]**  
**[PERCENTAGE]**

**Полезная информация**  
**10%**

**Ненужная информация**  
**15%**

**[CATEGORY NAME]**  
**[PERCENTAGE]**

**Информация не по теме**  
**15%**

**Смежная информация**  
**20%**

**Устаревшая по теме**  
**20%**

# Основные характеристики, которые отражают ИСд

- Системы классификации и кодирования информации и стандарты (УДК, ББК, классификаторы, MARC форматы и т.д.)
- Система формирования научных информационных ресурсов;
- Система поиска информации в локальных и корпоративных сетях;
- Базы данных научной информации (электронный каталог, полнотекстовые базы данных научных журналов, книг и т.д.);
- Система отображения информации (интерфейс, web pages и т.д.);
- Система администрирования баз данных научной информации;
- Создатели информационных ресурсов (авторы, библиотекари, операторы, специалисты по БД, администраторы баз данных и т.д.);
- Пользователь информационных ресурсов (юзеры);
- Телекоммуникационная инфраструктура (локальные сети, Интранет, Интернет).

# УРОВЕНЬ ИНТЕРЕСОВ ЧИТАТЕЛЕЙ РАЗЛИЧНЫХ ОБЛАСТЕЙ ДЕЯТЕЛЬНОСТИ

№	Категория интересов и материалы	Пользователь					
		Научный сотрудник	Преподаватель	Инженеры	Соискатели ученых степеней, магистранты и докторанты	Студент	Бизнесмен
		1	2	3	4	5	6
1	Научные исследования (статьи и книги)	98.3	82.5	40.2	99.4	20.1	12.6
2	Статьи по технологиям	22.4	32.1	99.2	70.3	56.6	42.2
3	Диссертации,	99.6	80	56.7	98.7	70.6	10.4
4	Материалы по проф. технологиям	51.4	50	99.6	50.4	47.3	21.3
5	Энциклопедии	92.4	63.9	51.3	54.3	42.4	11.2
6	Стандарты	10.6	32.5	90.4	41.5	21.5	52.4
7	Учебники, учебно-методические материалы	71.4	100	43.1	72.8	99.7	12.5
8	Статьи по бизнесу	22.5	20	47.1	31.6	13.1	99.8

# 10 критериев оценки информационного ресурса для поиска и его защиты

**A1 - начальная стоимость ресурса** (0 – бесплатный; 1- ниже среднего; 2- средняя; 3- выше среднего);

**A2 - ресурсная среда** (0 – отдельно на магнитных носителях или на компьютере; 1 - локальная сеть, 2 - корпоративная сеть, 3 - глобальная сеть);

**A3 - тип ресурса** (0 - общедоступная информация, 1 - корпоративная информация, 2 - платная информация, 3 - конфиденциальная информация);

**A4 - использование ресурсов** (0- очень редко; 1 редко; 2 – часто; 3 – очень часто);

**A5 - важность ресурса во всей информационной системе** (0 - сбой ресурса(его разрушение) не влияет на другие ресурсы; 1 - сбой ресурса приводит к остановке некоторых компонентов других ресурсов; 2 - сбой ресурса приводит к сбою нескольких компонентов других ресурсов; 3 - сбой ресурса приводит к сбоям в работе всей системы;

**A6 – затраты на восстановление ресурса** (0 - незначительные, 1 - средние, 2 - высокие, 3 – большие затраты или невозможные);

**A7 - время, затрачиваемое на восстановление ресурса** (0 - незначительное, 1 - среднее, 2 – требует много времени, 4 – требуется критически много времени);

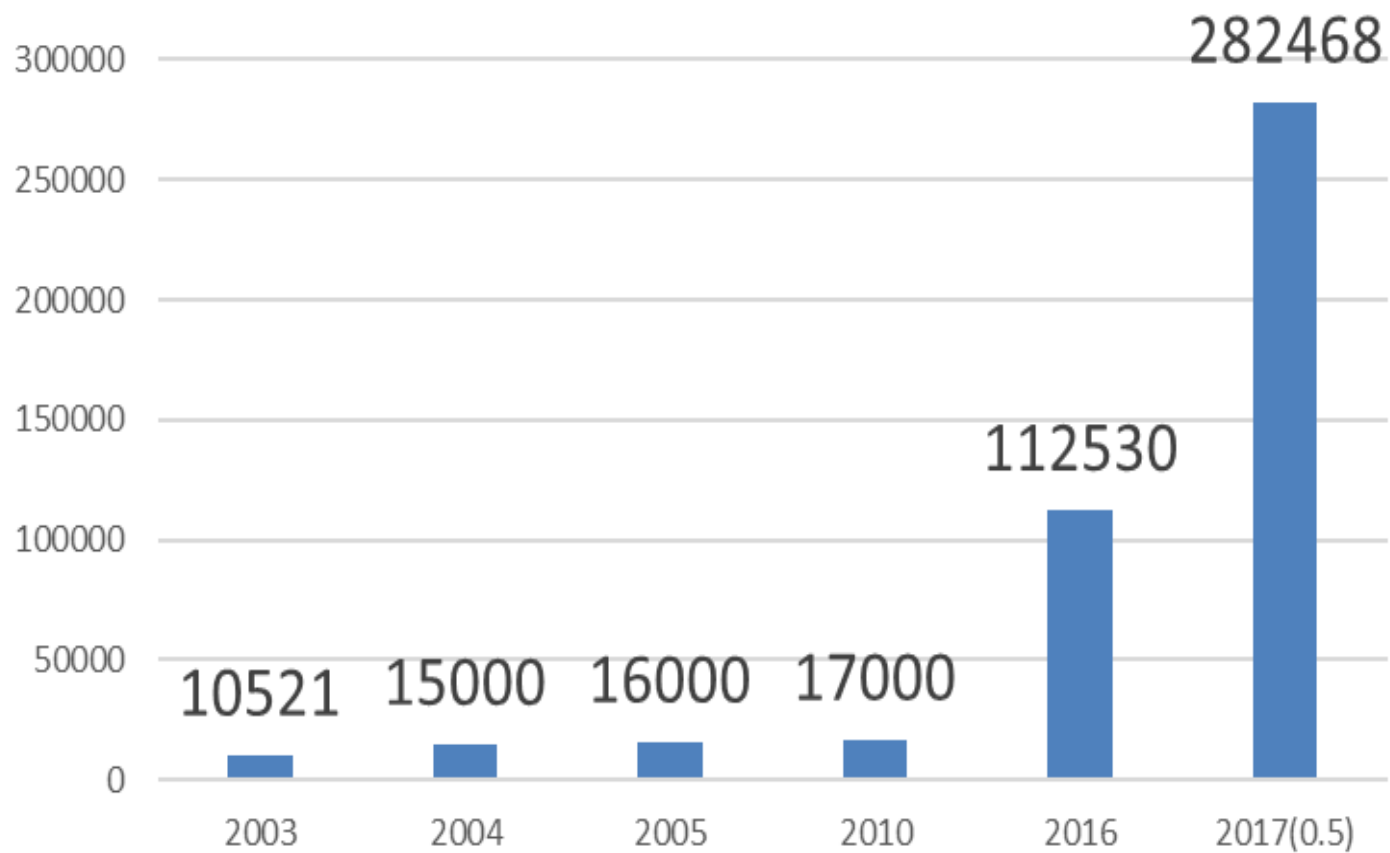
**A8 - возможность частично или полностью восстановить ресурс** (0 - легко, 1 - требует времени и средств; 2 - требует значительных временных и материальных затрат; 3 – невозможно восстановить);

**A9 - нарушение конфиденциальности** (0 - не наносит существенного ущерба; 1 - наносит материальный ущерб в определенных обстоятельствах; 2 - приводит к меньшему моральному и / или материальному ущербу; 3 - причиняет значительный ущерб или отказ всей системы);

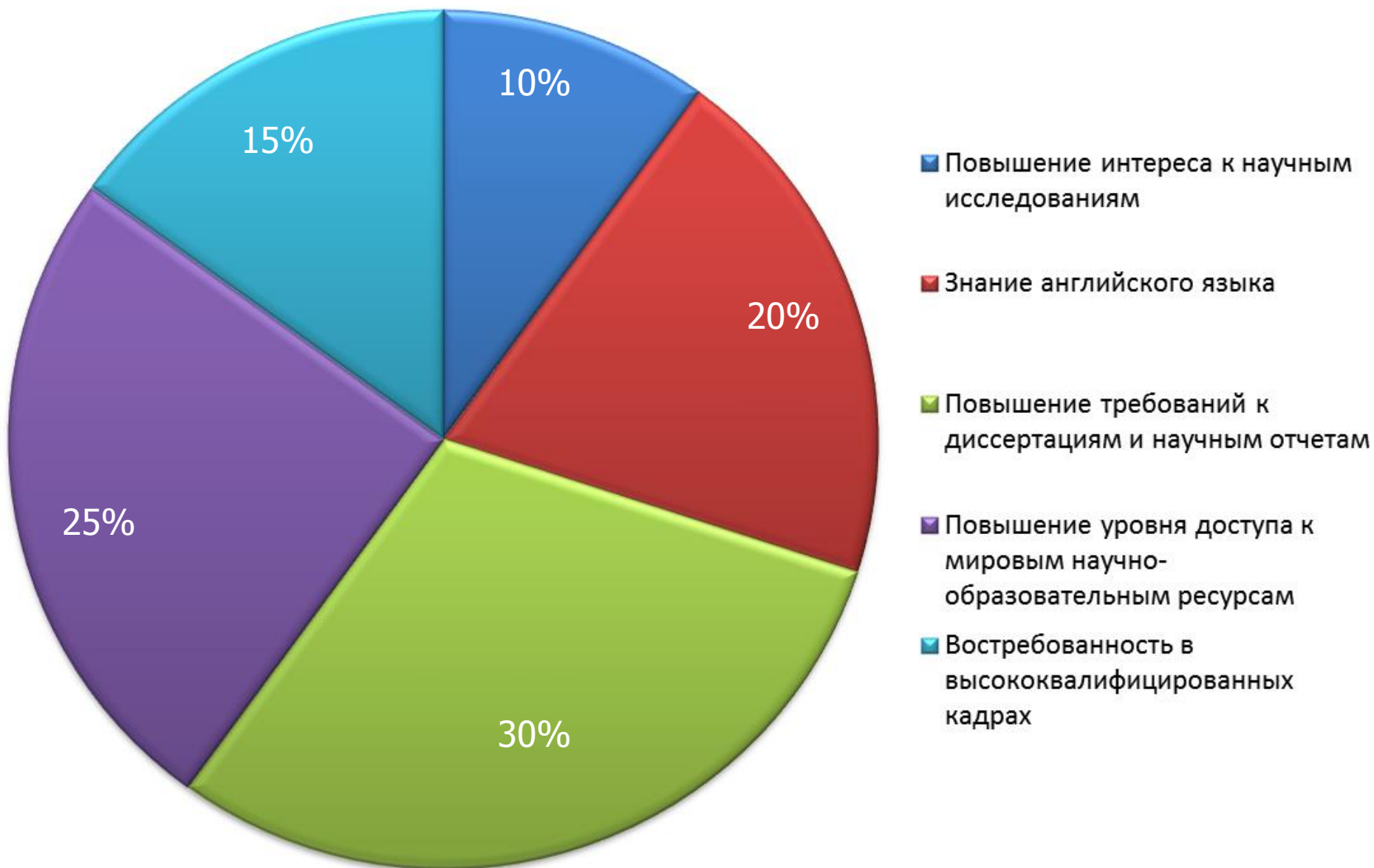
**A10 - нарушение целостности данных ресурса** (0 - не приводит серьезным последствиям; 1 – результаты заметны(ощутимы), но не приводят к прекращению работы; 2 - Результатом будет неисправность; 3 - следствием этого является искажение данных ресурсы и невозможность их исправить).



## Динамика использования баз данных EBSCO Information Services



# ВЛИЯНИЕ ФАКТОРОВ АКТИВИЗАЦИИ ИСПОЛЬЗОВАНИЯ ЦЕННЫХ ЭЛЕКТРОННЫХ НАУЧНЫХ ИСТОЧНИКОВ



## Инфлюэнты публикационной активности

1. 28% - Сам стимул написания статьи для научных сотрудников, преподавателей, докторантов. Учет публикаций в высокорейтинговых журналах при аттестации, влияние на карьерный рост;
2. 27% - Наличие ценного научного материала, научный потенциал. Уровень научных исследований в организации, наличие лабораторий с современным оборудованием и др.
3. 25% - Навыки написания научных статей. Умение оформить статью под формат тех престижных журналов под которые подходит тема статьи.
4. 15% - Знание иностранных языков. Главным образом – английского языка, который является языком общения научных сообществ.
5. 5% - Информированность о научных журналах. В настоящее время этот показатель все меньше и меньше влияет на публикационную активность, т.к. списки научных журналов можно легко найти в Интернете, в списках Web of Science и Scopus.

## **Инфлюентный анализ для выявления наиболее значимых показателей для поиска информации в корпоративных библиотечных сетях включает следующие этапы:**

- Систематизация и классификация информации в базах данных корпоративной информационно-библиотечной сети;
- Проведение экспертного опроса или других способов на предмет выявления наиболее существенных факторов, влияющих на поиск данных.
- Инфлюентный анализ факторов. Выявление параметров, которые в той или иной мере влияют на процесс формирования и поиска данных;
- Внесение изменений в процедуры (алгоритмы и программы поиска) поиска данных с учетом инфлюентных факторов;
- Возможность итерации при систематизации и оценке ресурсов для повышения эффективности поиска.

# Использование методов инфлюентного анализа позволяют

- существенно сократить объем поиска за счет отсекаания массивов данных неадекватных запросу.
- выявить наиболее важные показатели, на которые нужно уделять наибольшее внимания, выделять средства для оптимизации суммарного показателя эффективности;
- снизить затраты на те факторы, которые в настоящее время не имеют существенного значения для нахождения нужной информации;
- управлять некоторыми факторами, чтобы улучшить работу всей системы;
- снизить влияние тех показателей, на которые мы в данный момент не можем повлиять (управлять).

*Спасибо за внимание !*

*marat56@mail.ru*

*+998946883133*

