

ИТ в юриспруденции



Лекция 6. Информационный рынок: онлайновые СМИ и справочно-правовые системы

Розина Ирина Николаевна
Д.п.н., профессор кафедры ИТ

Основные определения

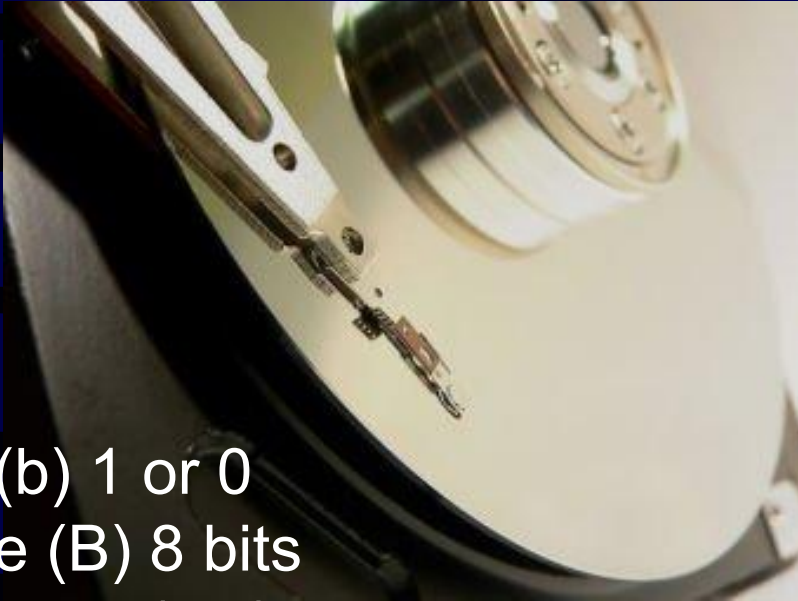
- **Рынок информационных услуг** – совокупность экономических, правовых и информационных отношений по торговле (продаже / покупке) между поставщиками (продавцами) и потребителями (покупателями), характеризуется определенной номенклатурой услуг, условиями и механизмами их представления и ценами
- **Информация** – сведения о лицах, предметах, фактах, событиях, явлениях и процессах независимо от формы их представления. Является товаром на рынке информационных услуг
- **Информационные ресурсы** – информация в виде отдельных документов и массивов документов, зафиксированные на материальных носителях и хранящиеся в АИС (библиотеках, архивах, фондах, банках данных и пр.)

Рынок информационных ресурсов

- **Деловая (бизнес) информация**
- **Научно-техническая** (*документальная, фактографическая, библиографическая, реферативная и полнотекстовая информация в фундаментальных и прикладных исследованиях*) и **специальная профессиональная информация** (*ресурсы для юристов, врачей, инженеров и т.д.*)
- **Массовая потребительская информация** (*новости, справочная, развлекательная и справочная информация – погода, расписания, покупки и продажа, справочники различных служб и пр.*)



Цифровые устройства, 2006



Bit (b) 1 or 0

Byte (B) 8 bits

Kilobyte (KB) 1,000 bytes

Megabyte (MB) 1,000 KB

Gigabyte (GB) 1,000 MB

Terabyte (TB) 1,000, GB

Petabyte (PB) 1,000 TB

Exabyte (EB) 1,000 PB

Zettabyte (ZB) 1,000 EB

Yottabyte (YB) 1,000 ZB

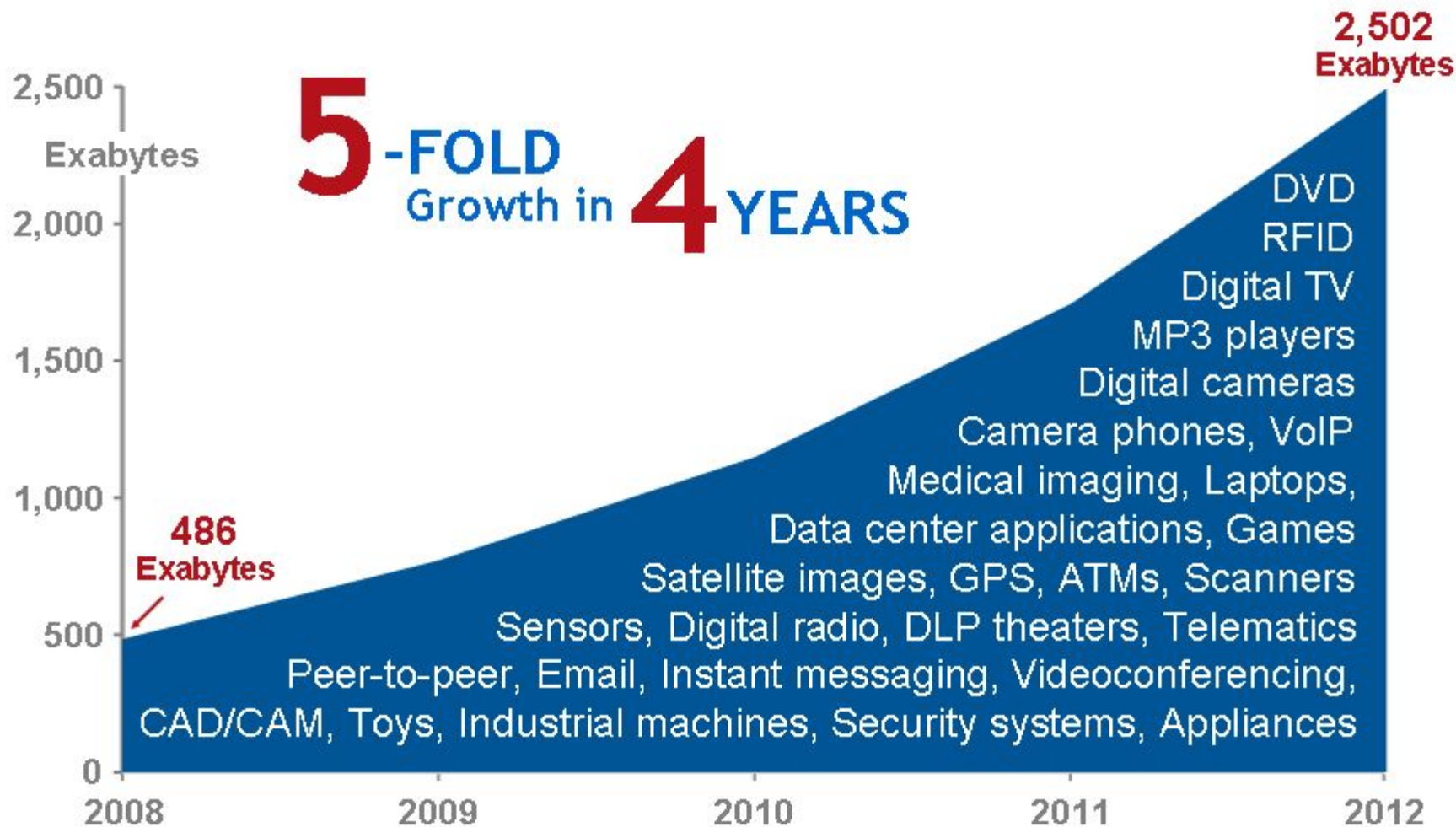
- **2002** — объем цифровой информации превысил бумажный

Устройства Число (млн)

- Digital Cameras 400
- Camera Phones 600
- PCs 900
- Audio Players 550
- Mobile Subscribers 1600
- LCD/Plasma TVs 70

Digital Information Created and Replicated Worldwide

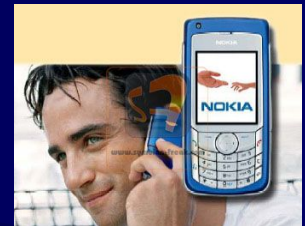
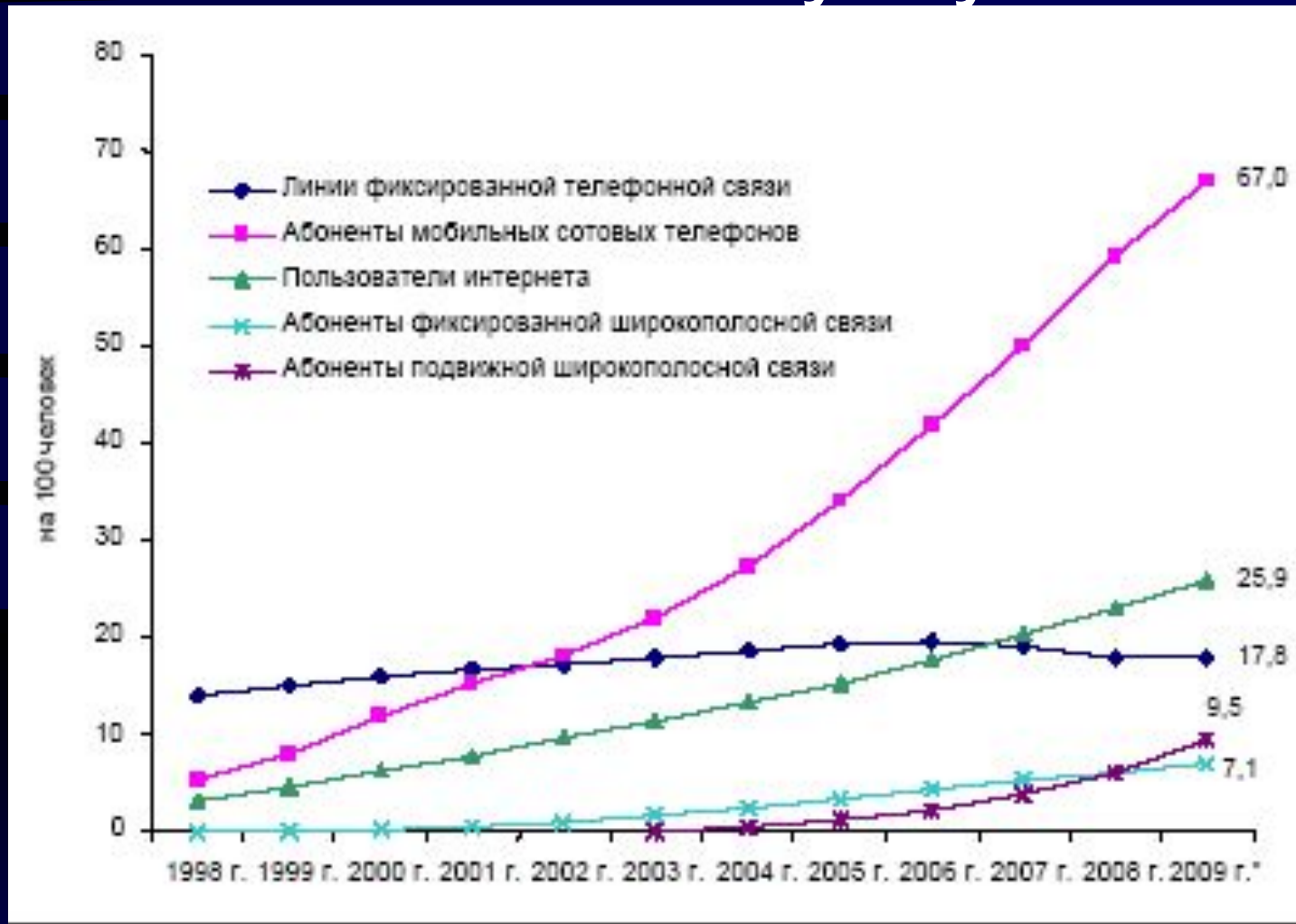
EMC²
where information lives[®]



Source: IDC Digital Universe White Paper, Sponsored by EMC, May 2009

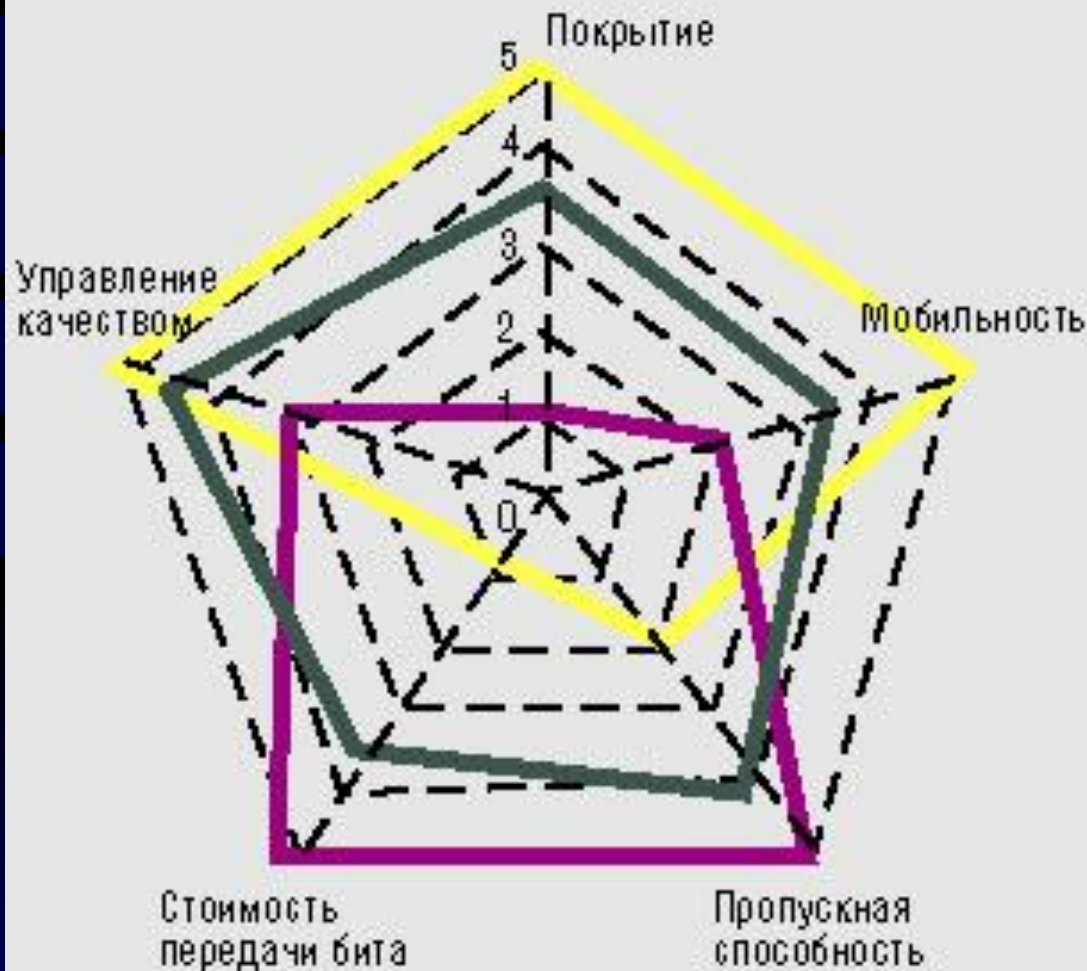
© Copyright 2009 EMC Corporation. All rights reserved.

Усредненные данные использования услуг ИКТ в мире



Источник: Всемирная база данных МСЭ по показателям в области электросвязи /ИКТ

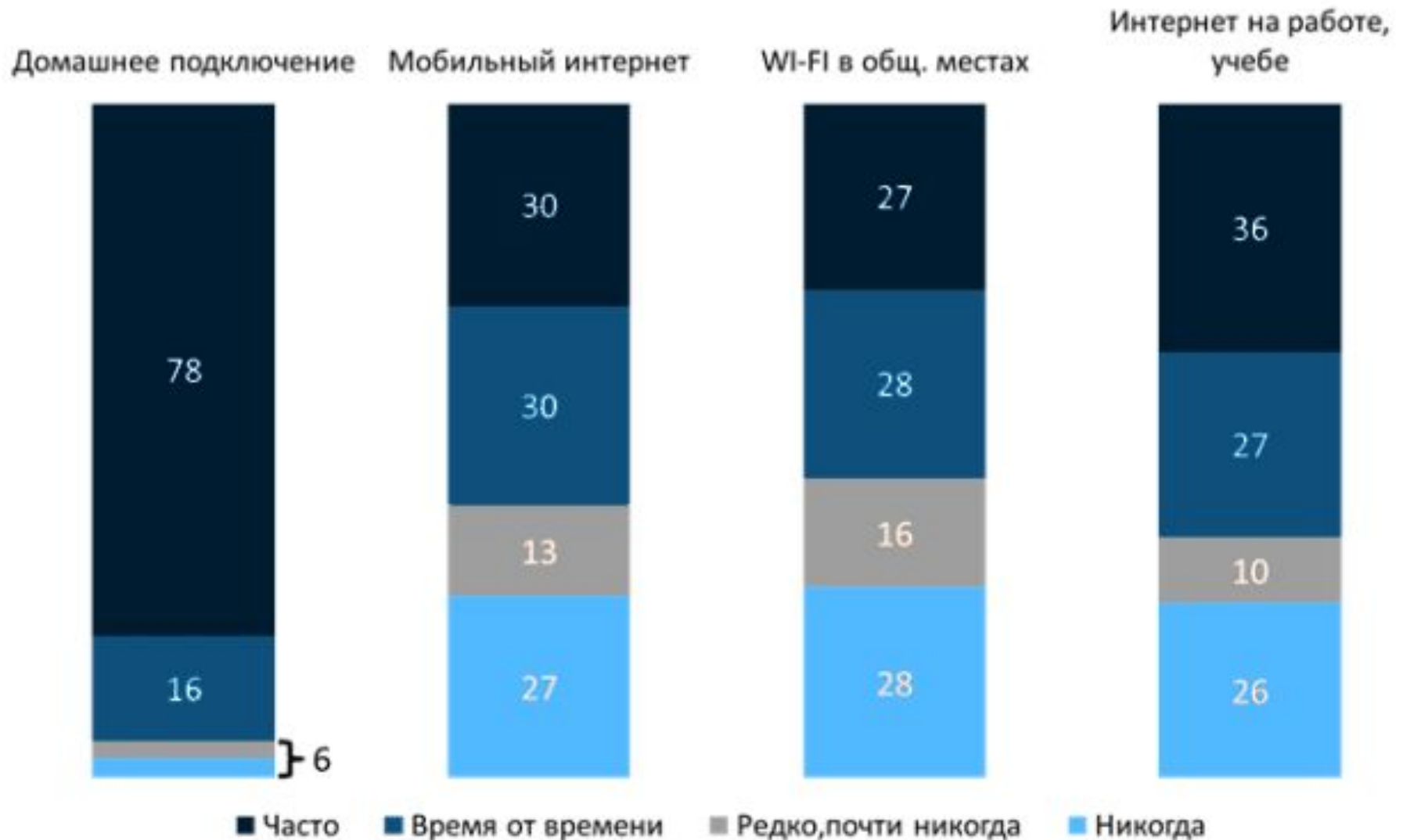
Качественное сравнение технологий беспроводного доступа WiMAX, WLAN, сотовая связь



Источник: Samsung и Unstrung Insider.

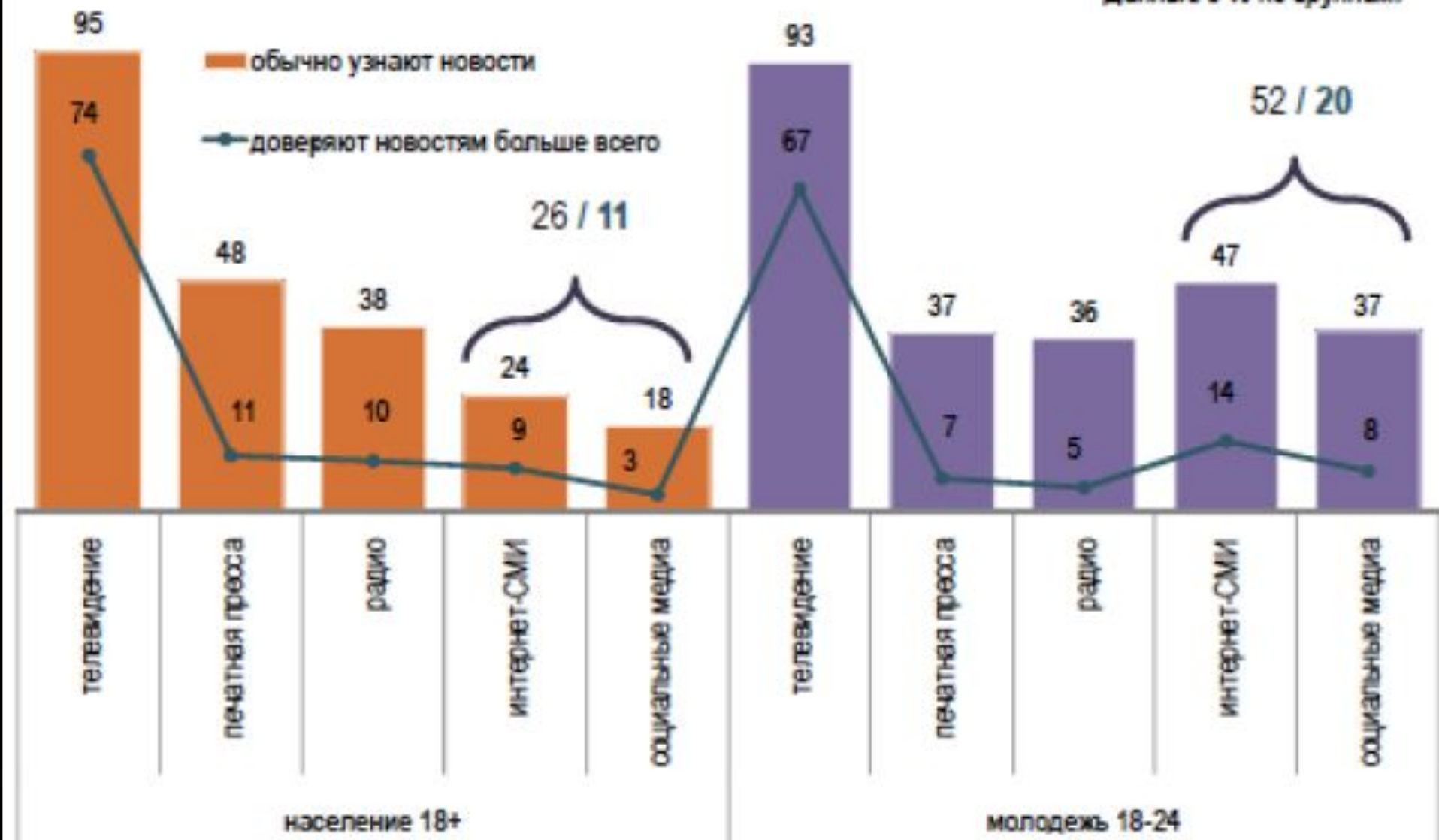
Рис. 1. Качественные сравнения технологий беспроводного доступа — мобильного WiMAX, WLAN, сотовой связи

Как часто для выхода в Интернет Вы используете... (в %, Москва, март 2014)

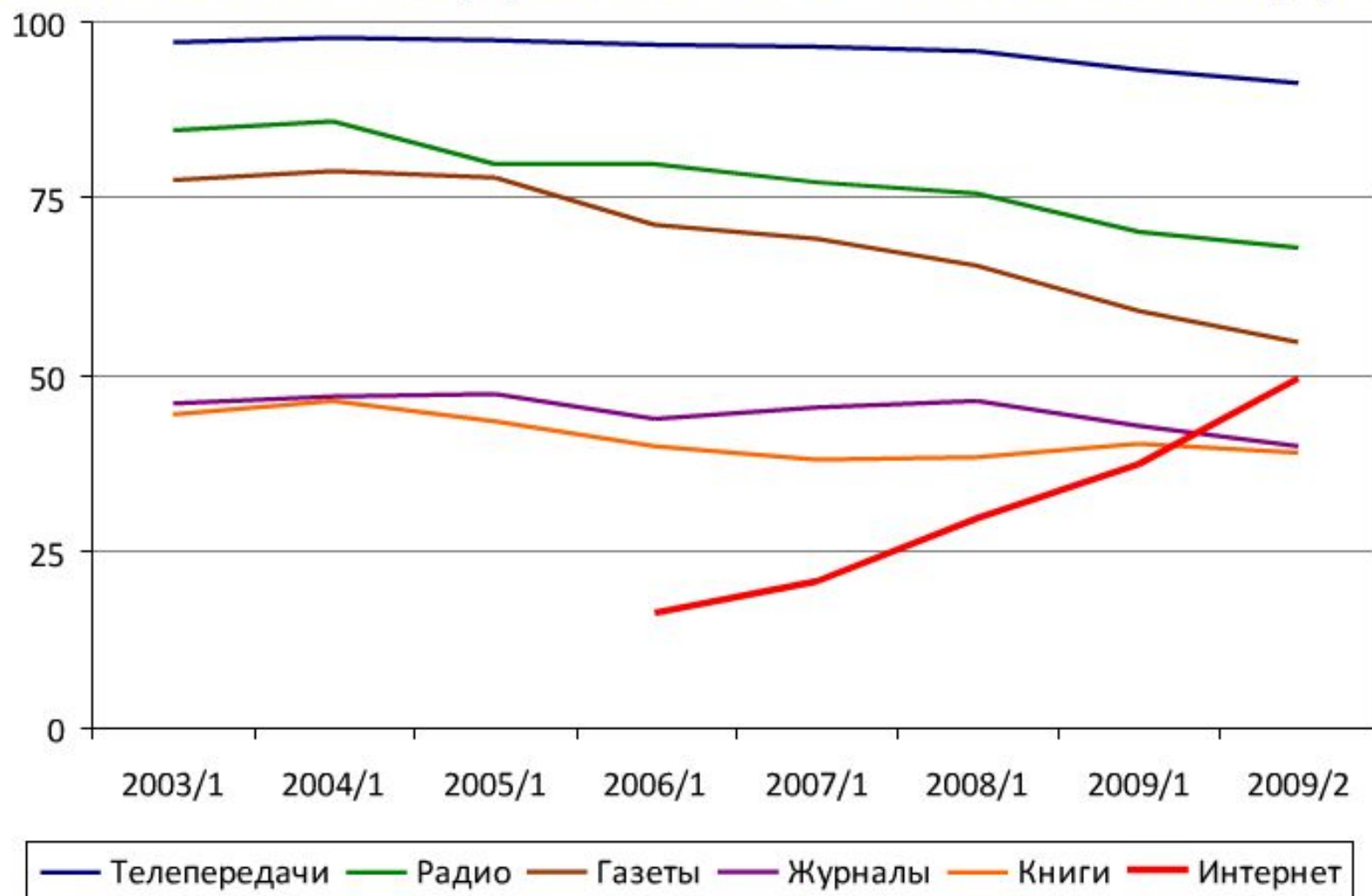


Источники новостей для населения России (ФОМ, 2011)

Данные в % по группам.



Изменение недельного охвата медиа



- Недельный охват Интернета превысил охват журналов и книг. Продолжает сокращаться недельная аудитория газет и радио.

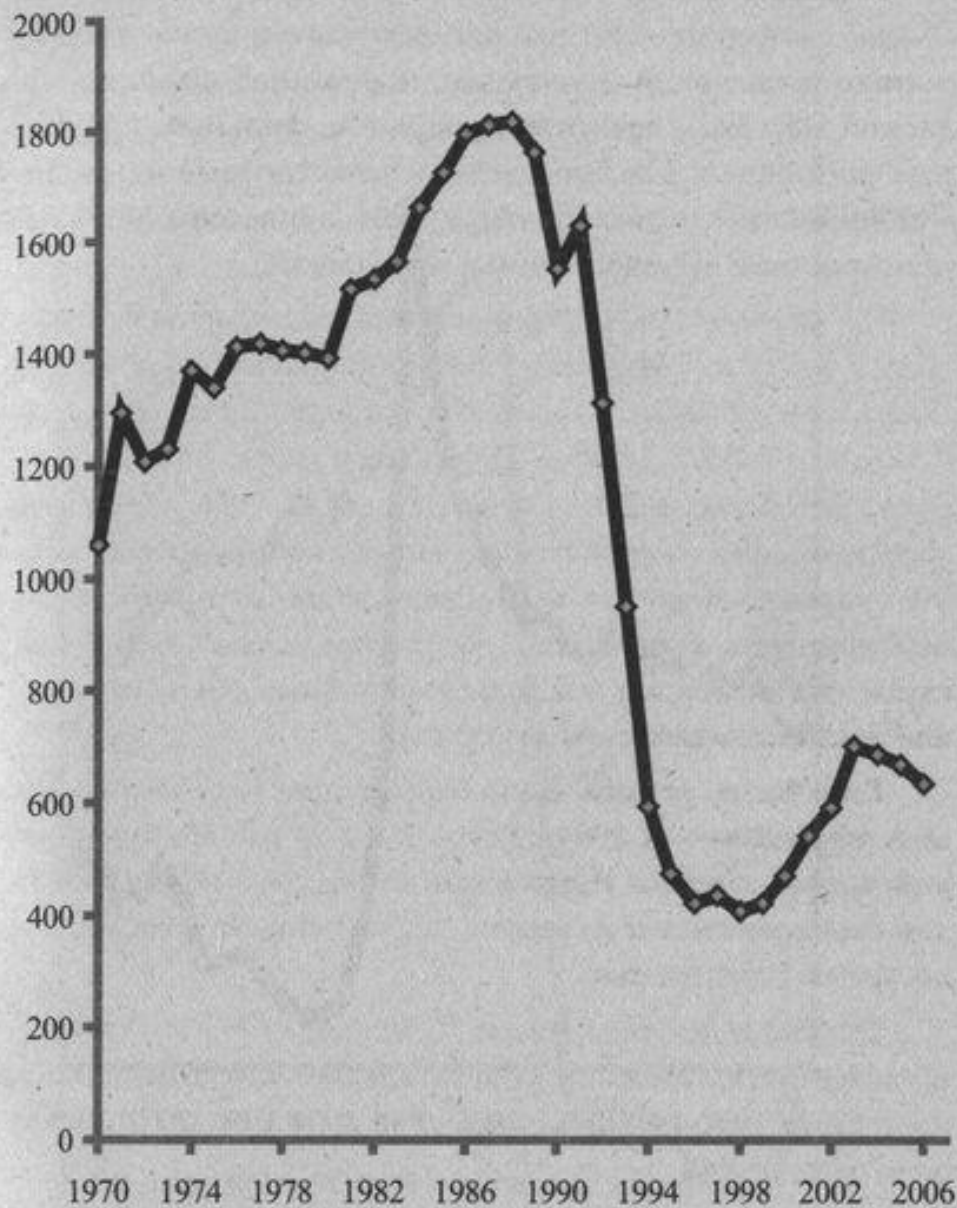


Рис. 3-32. Тираж изданных книг и брошюр
в РСФСР и РФ, млн. экз.

Тираж изданных
книг и брошюр в
РСФСР и РФ, млн.
экземпляров, в
1970-2006 годах

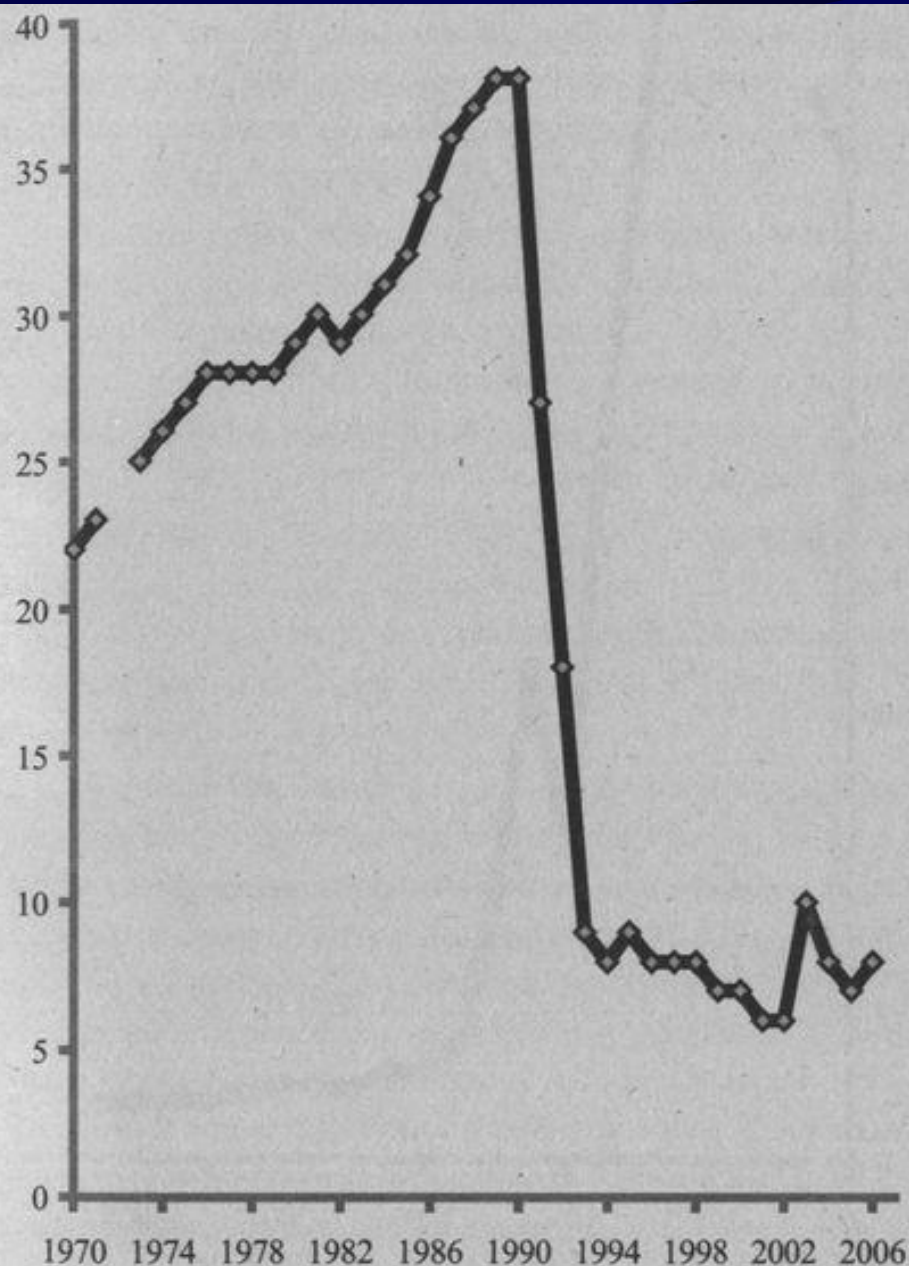
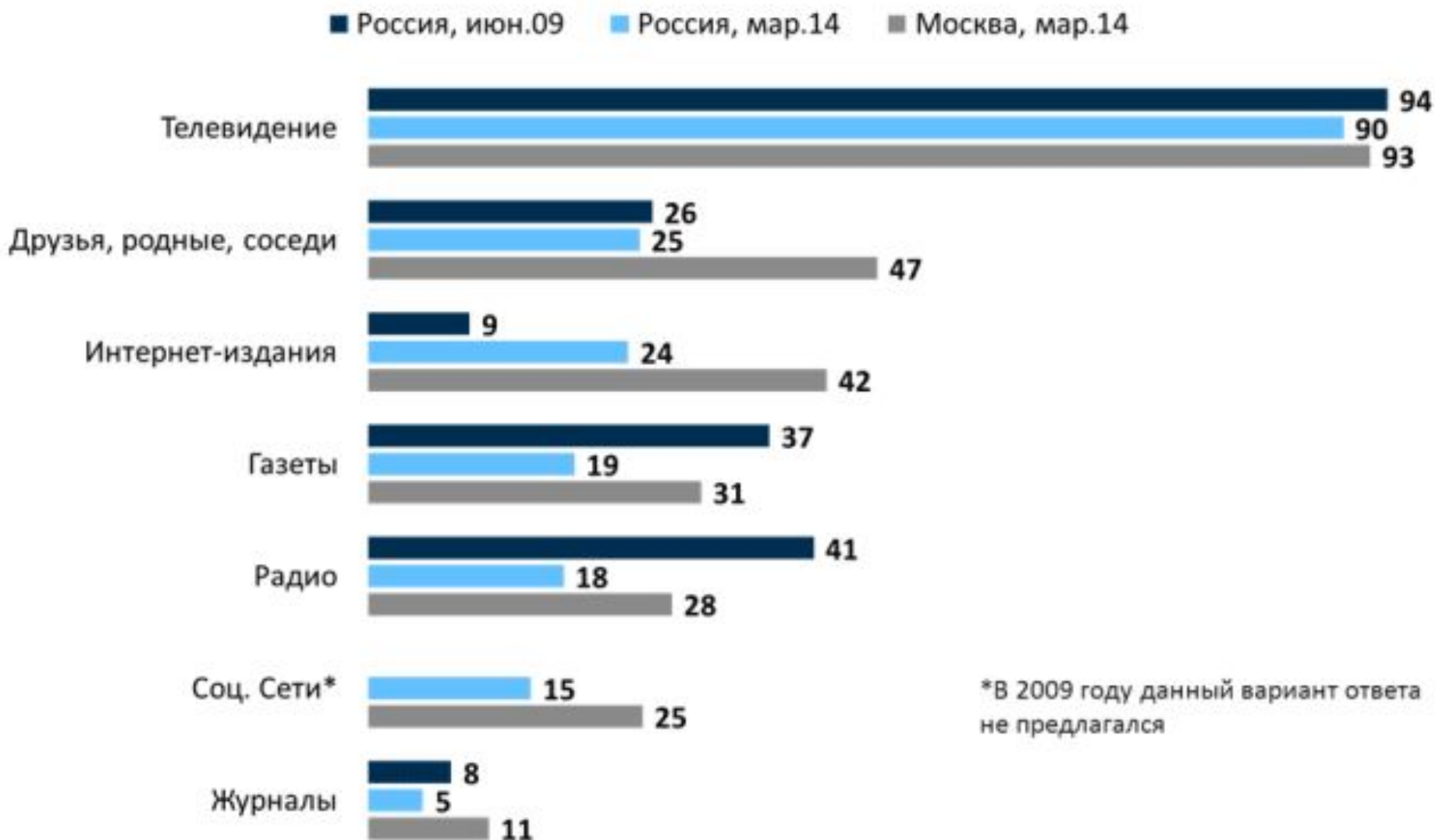


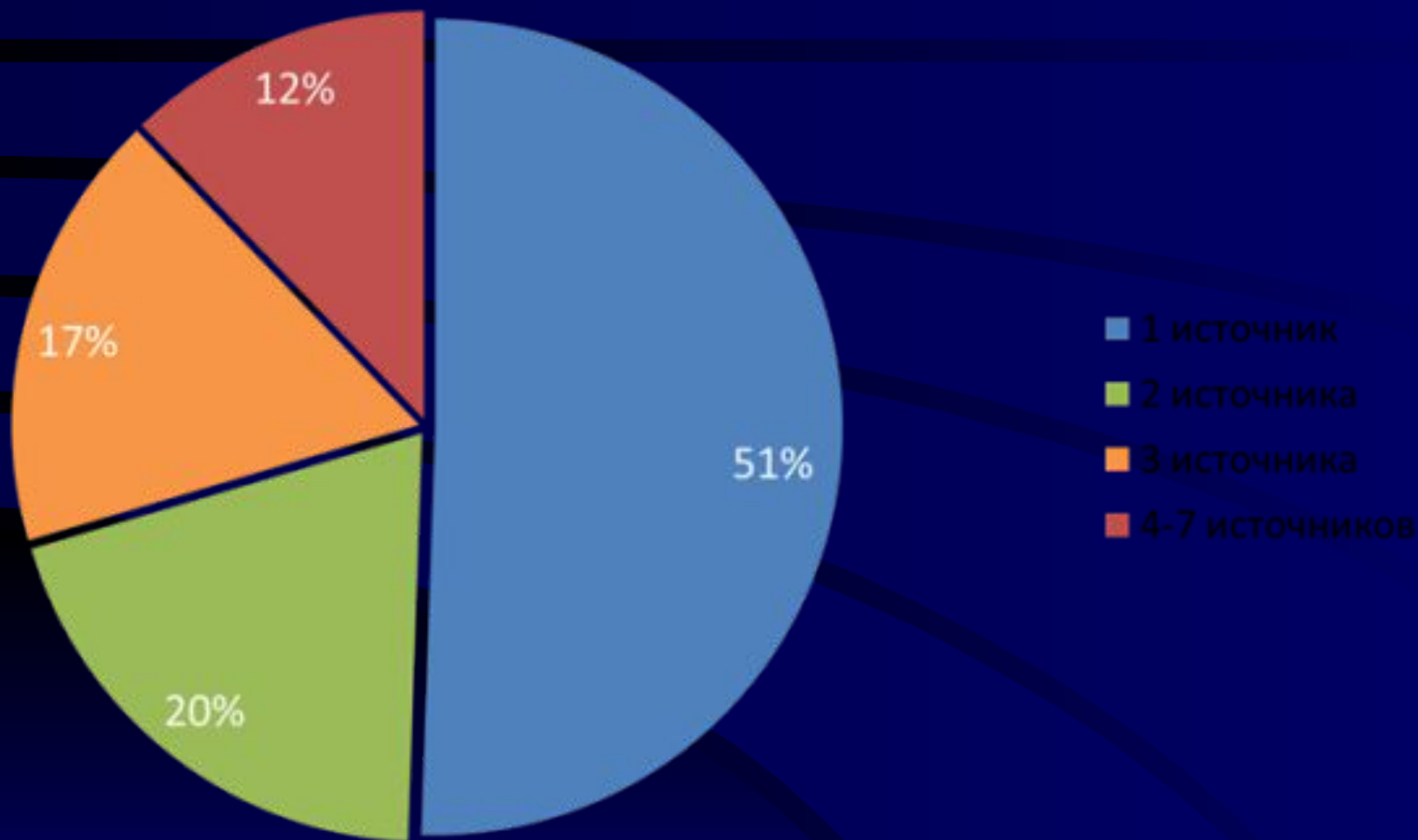
Рис. 3-34. Годовой тираж газет в РСФСР и РФ, млрд. экз.

Годовой тираж газет
в РСФСР и РФ в
1970-2006 годах,
млрд. экземпляров

Откуда Вы чаще всего узнаете о новостях в стране и в мире? (в %)

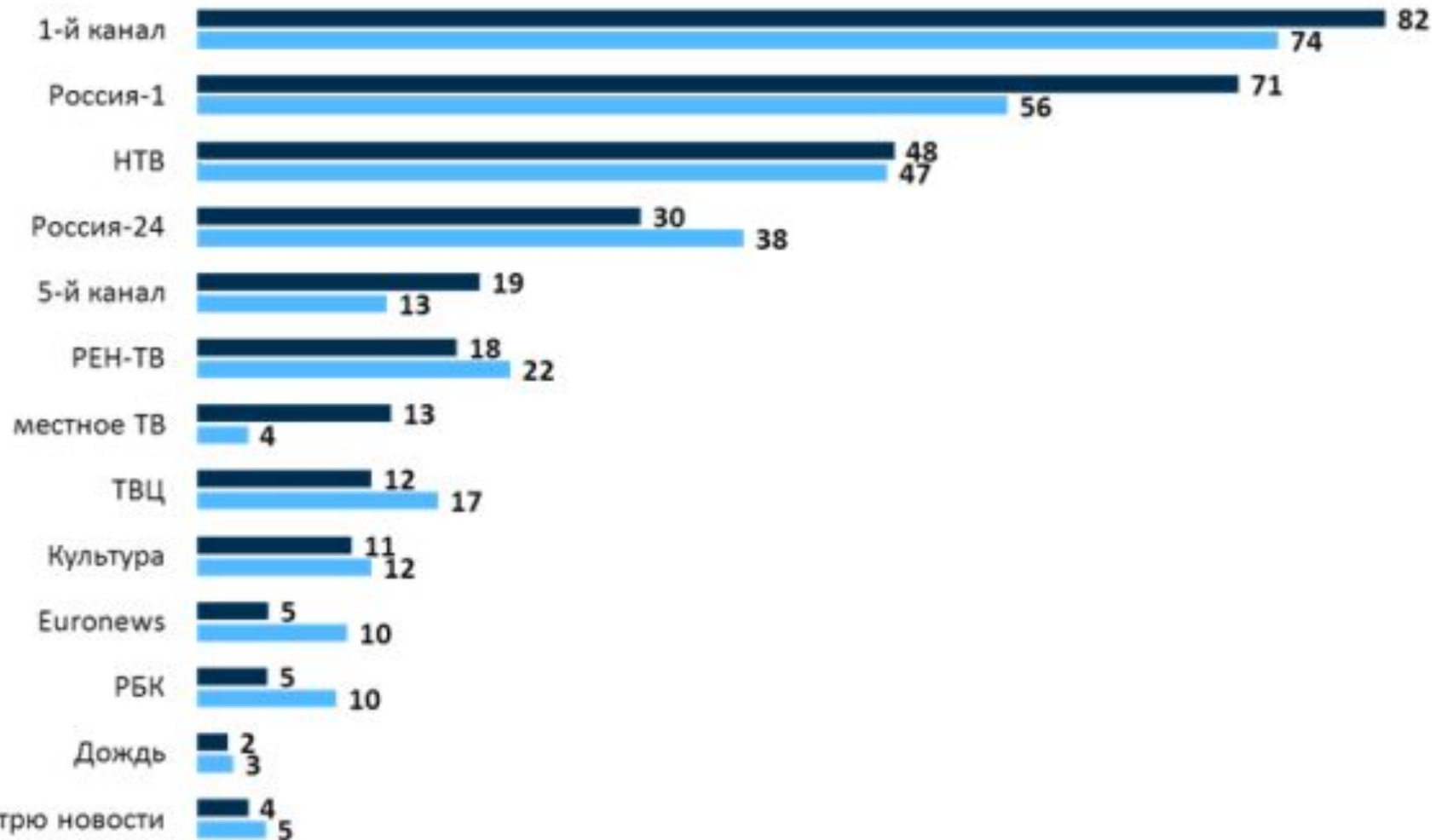


Получение новостей из нескольких источников (в %, Россия, март 2014)



Смотрите ли Вы новости по телевизору, и если да, новости каких каналов Вы смотрите более или менее регулярно? (в %)

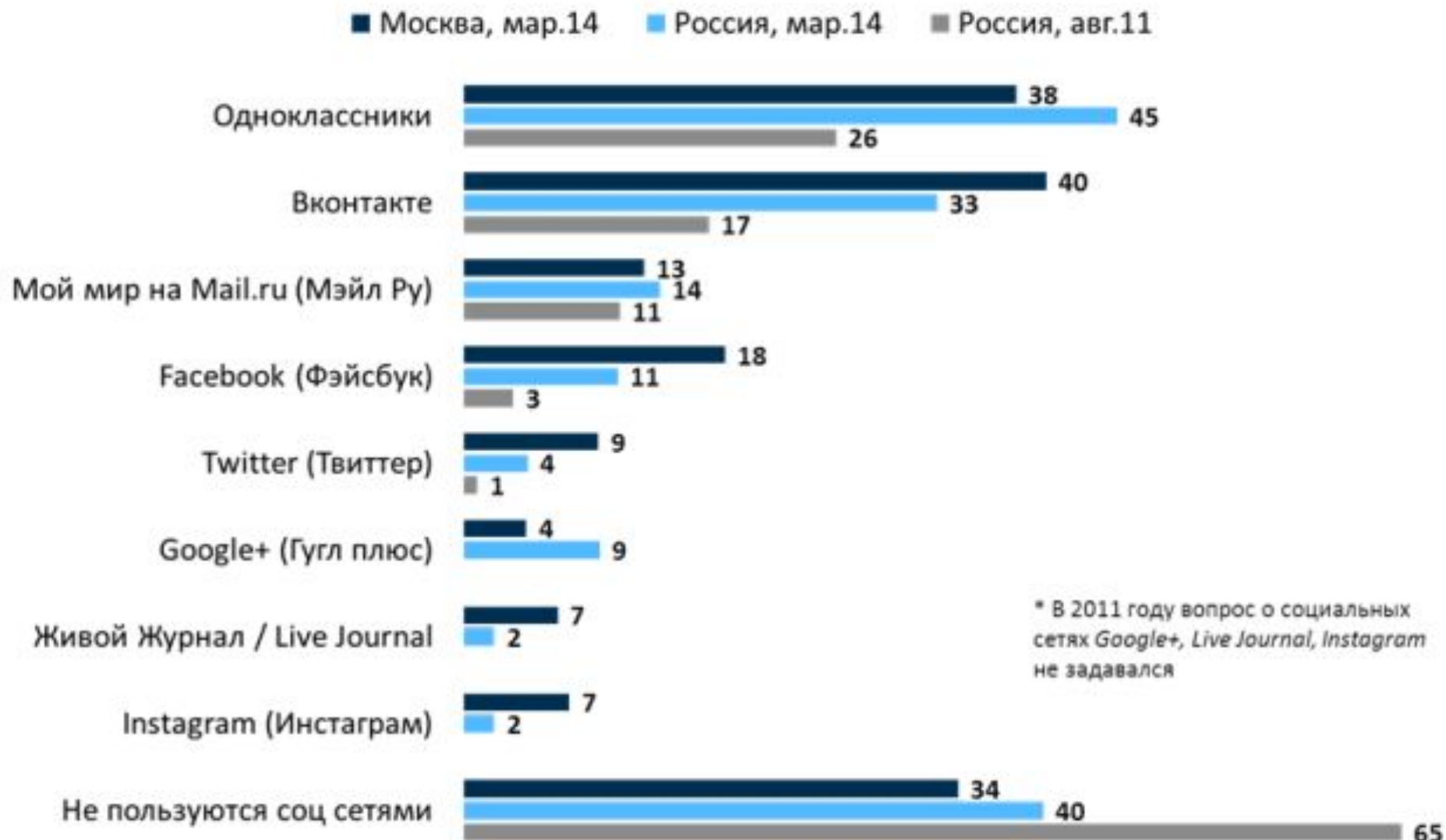
■ Россия (мар.14) ■ Москва (мар.14)



Цели пользования Интернетом (в %, Россия, май 2014)

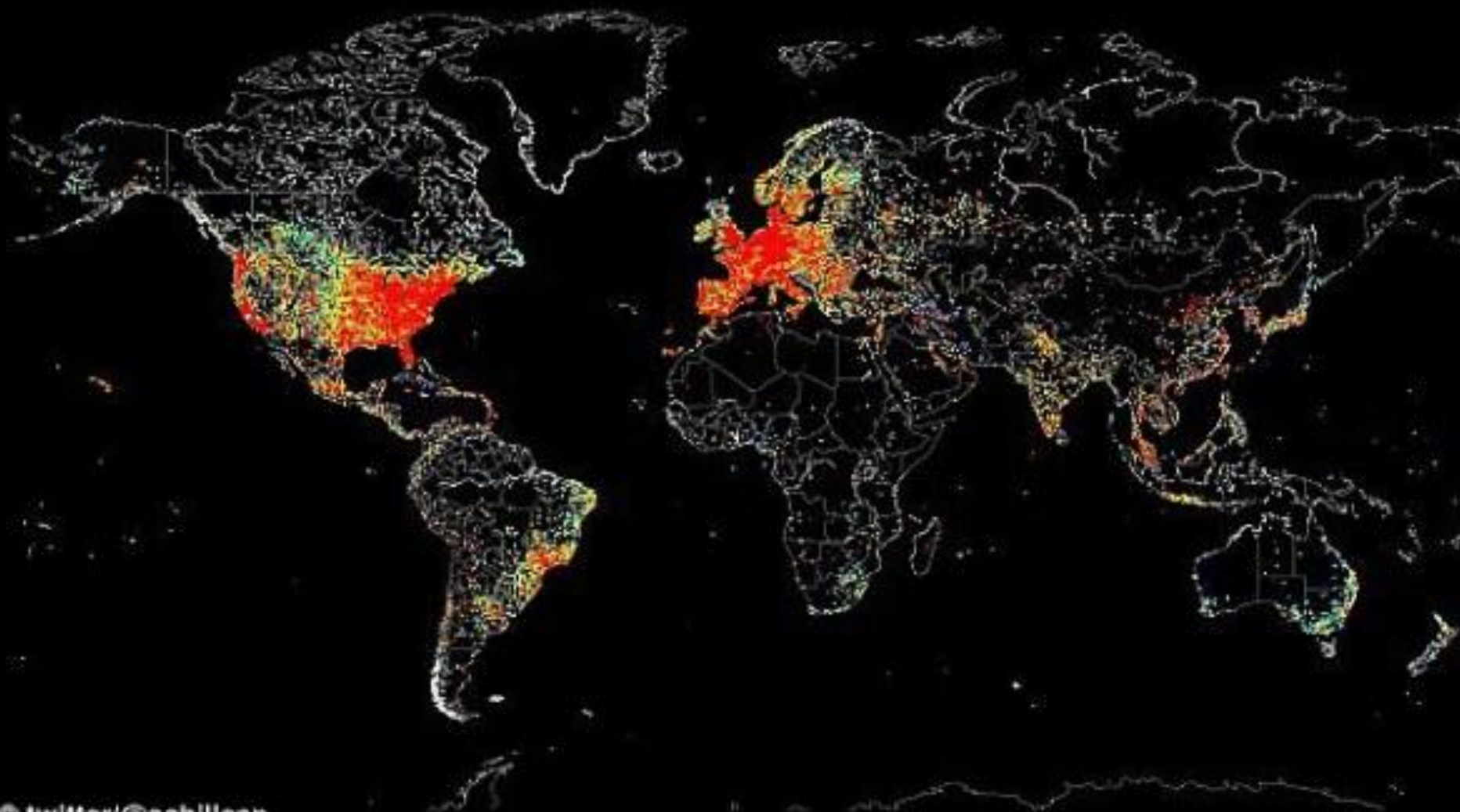


Посещение различных социальных сетей (в %)

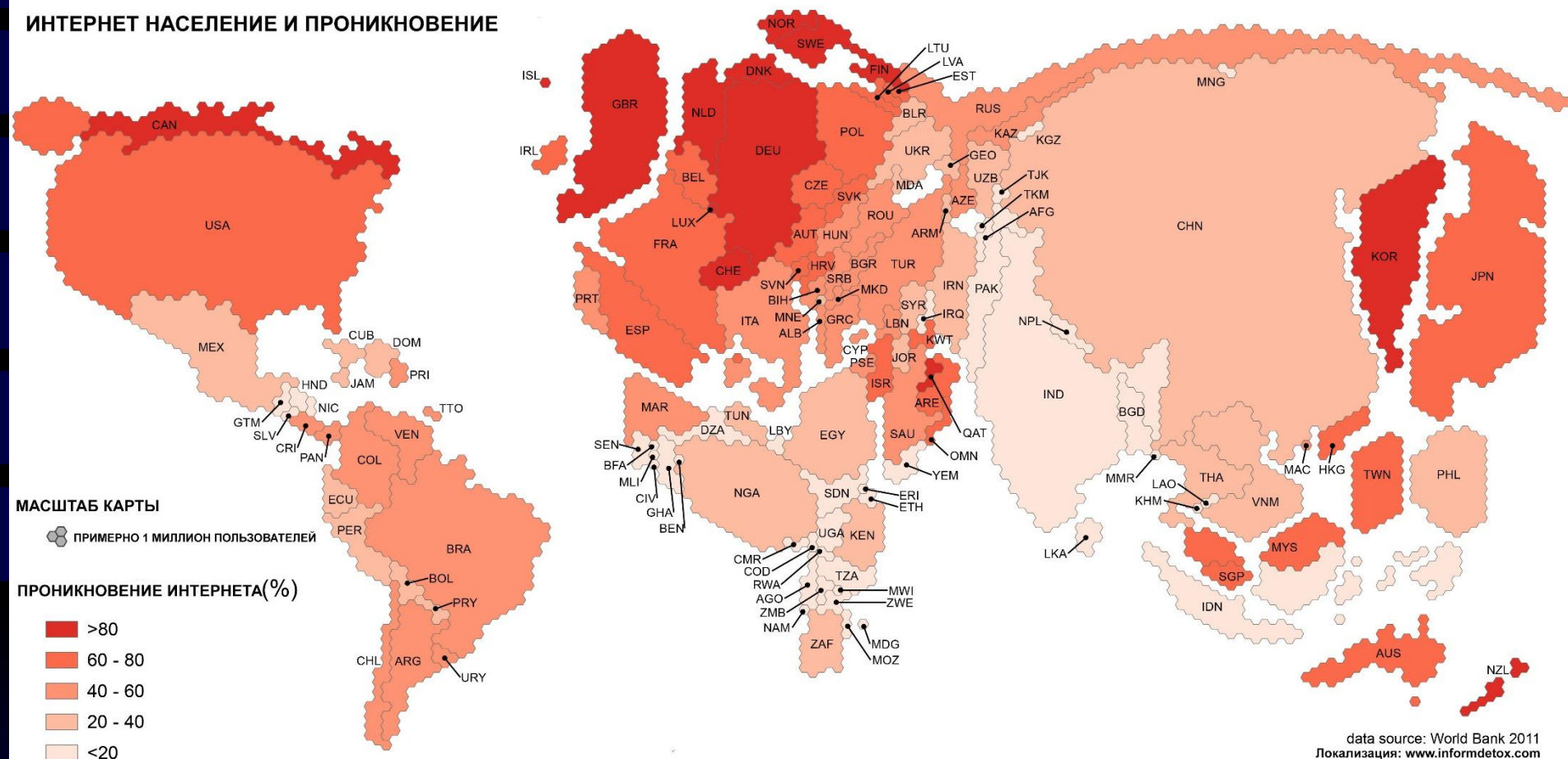


Плотность веб-соединений в мире

SHODAN



ИНТЕРНЕТ НАСЕЛЕНИЕ И ПРОНИКНОВЕНИЕ

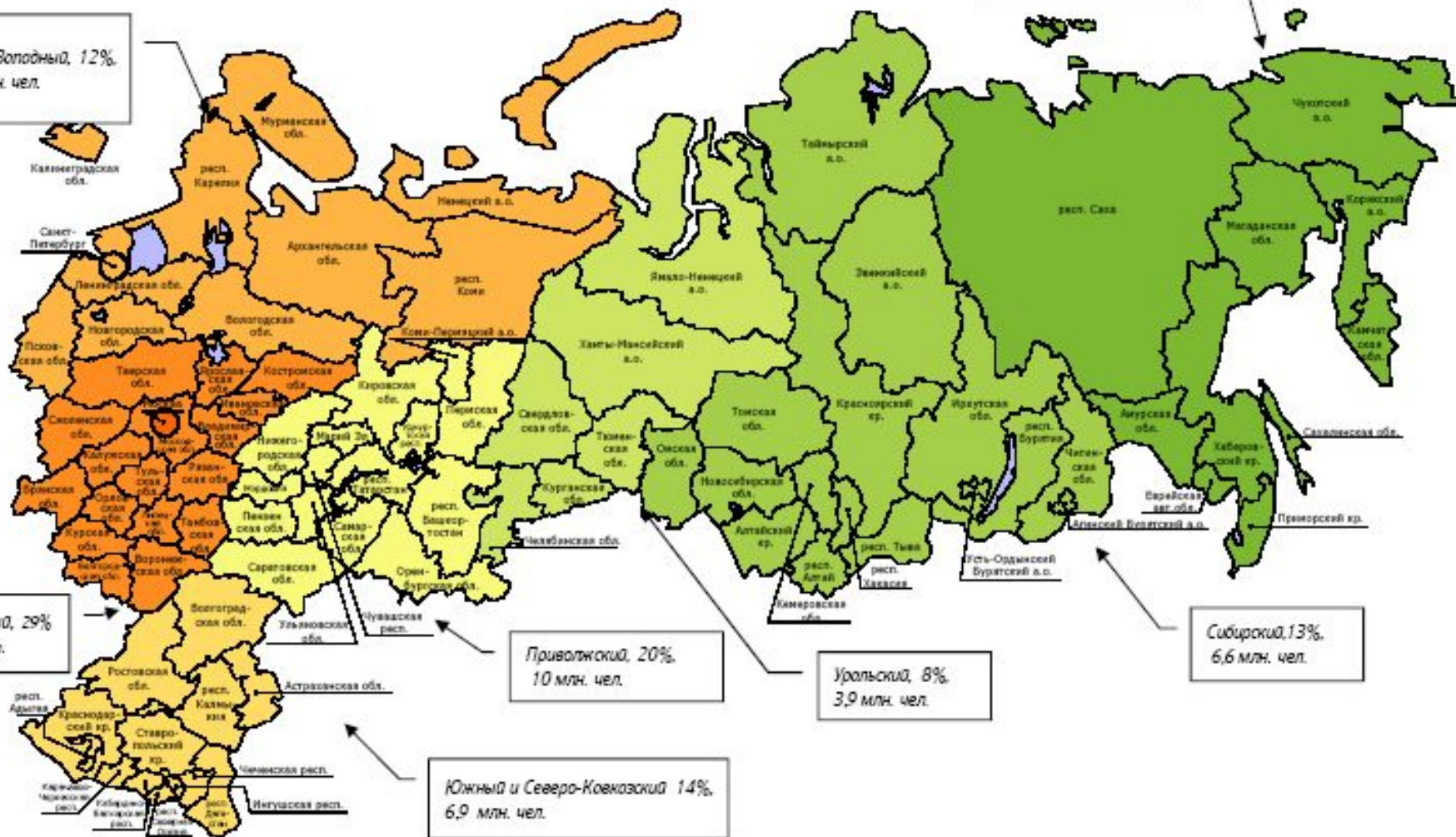


Распределение общего числа пользователей Интернета по регионам

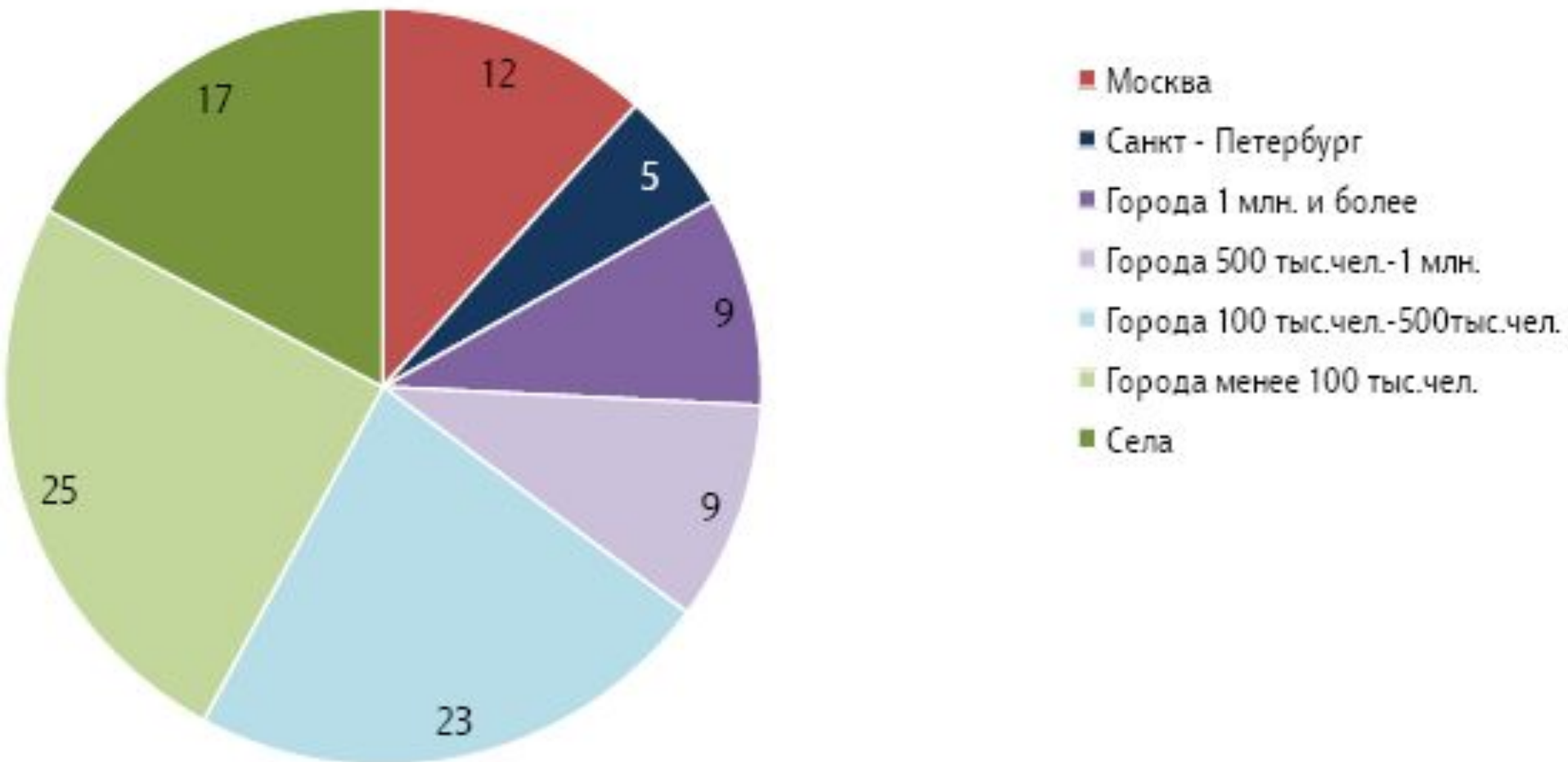
Россия в целом: 100% (50,3 млн. чел.)

Северо-Западный, 12%,
6,1 млн. чел.

Дальневосточный, 5%,
2,4 млн. чел.



Структура пользователей Интернета по типам населенного пункта, % от пользователей



Сетевая активность пользователей Рунета



Топ-15 интернет-сайтов

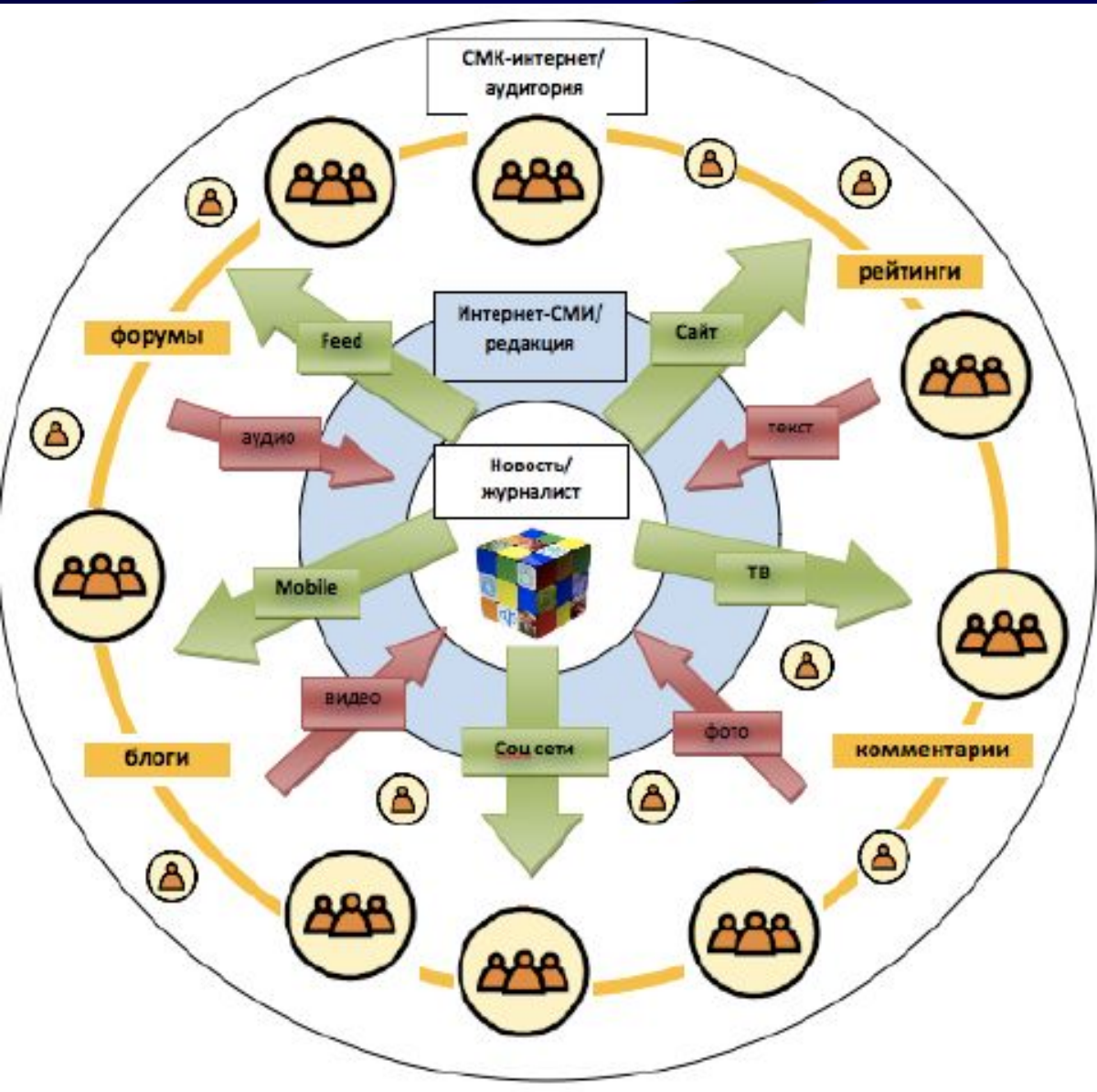
В среднем за месяц



В среднем за день



Количество
посетителей,
млн чел., ФОМ,
2010



Интерактивные
конвергированные
социальные медиа

Электронный журнал

electronic, online, web journal, magaozine

Публикации, сохраняющие большинство характеристик печатных изданий, но которые производятся и распространяются онлайн

- **Середина 1990-х годов** – электронные версии коммерческих изданий научных журналов на CD-ROM
- Научные **электронные журналы** по проектам “Mercury”, CORE, “Tulip”, 1987–1993; JSTORE и “High Wire Press”, с 1995
 - существуют *только как электронные*, в онлайн-версии, нет печатной версии, публикуются по мере поступления статей (e-zine, webzine)
 - электронные *версии обычных журналов*, являются вторичными по отношению к печатной версии
 - *рефераты статей в открытом доступе* (журналы серии “Открытые системы”) и *полнотекстовые статьи* в закрытом с (для подписчиков, членов профессионального сообщества, авторов)

THE DIGITAL LABORATORY BOOKS WORKFLOW

The Digital Lab is where we preserve knowledge, making it available in different media other than books

Books Arrival in batches from the library

1 Check-In

Books meta-data is entered in repository database



Check-In Module

2 Scanning

Books are scanned at a resolution of up to 600 dpi



3 Processing

Image editing: refining quality by removing noise and unifying size

LAUDE LEV

PROCESSING ERRORS

SCANNING ERRORS

4 QA-Processing

Verification of digitization and meta-data



5 OCR

Converting image to searchable text



OCR ERRORS

ENCODING ERRORS

6 Encoding

Image on text as PDF and DjVu



7 QA-PDF

Verification of previous process



10 Check-Out

Process finished. Books are checked out of the Digital Lab.

Data Retrieval

9 Publishing

Files published on DAR:
<http://dar.bibalex.org>



8 Backup and Archiving

Files are saved in an online storage (PetaBox) & offline storage in DVDs



Check-Out

Ограничения возможностей переработки информации человеком

- *Физиологический барьер*
- Физиологические возможности человека ≈ 50 бит/сек
- *Технический барьер* (доступность, качество и надежность систем, проблемы поиска)
- *Психологический барьер* (ослабление внимания, интеллектуальной восприимчивости, мозаичность индивидуальной памяти)
- *Социальные, культурные, политические, языковые, географические, исторические и др. барьеры*

Физиологический барьер

*Информационная перегрузка (information overload),
информационный стресс, стрессы повторяющихся
действий, зрительный синдром*

- *Позитивная избыточность* – распространение дополнительной информации, не имеющей прямого отношения к проблеме, но призванной повысить эффективность восприятия основной информации
- *Негативная избыточность* – нарушение информационного процесса, наличие препятствий восприятия основной информации (“шум” или “помехи”)

Технологический барьер

- Трафик Интернета ↑ ежемесячно на 10%
- 1000÷1500 хостов ежемесячно
- «Период полураспада» информационных ресурсов ≈ 4 года
- За год любой указатель (справочник) сетевых ресурсов устаревает $\approx$ на $1/4$
- ≈ 27 тыс новостей в будний день
- 6 тыс новостей в выходной
- Дублирование новостей $\approx 20\%$



Психологические последствия работы в Интернете

- Пополнение знаний
- Чувство включенности в глобальное сообщество
- Развитие личности
- Потеря времени
- Увлечение в ущерб другим делам
- Утрата реального общения



Интернет-аддикция



- Впервые исследована в 1994 г.

К. Янг, Питтсбургский университет

- Термин *Internet Addiction Disorder* впервые введен И. Голдбергом в 1995 г. в качестве пародии (набор диагностических критериев для определения заболевания, но без клинических материалов, подтверждающих обоснованность введения нового типа заболевания)
- Аналогии: филателия, музыка, спорт, киномания, туризм, бега, аквариумные рыбки, азартные, компьютерные игры игроголики, работоголики

Интернет-аддикция

Симптомы:

- Неспособность, нежелание отвлечься
- Досада и раздражение при вынужденном отвлечении
- Навязчивые размышления об Интернете
- Побуждение тратить больше времени и денег
- Обман близких и приуменьшение длительности и частоты работы
- Склонность забывать о других делах
- Эмоциональный подъем
- Пренебрежение здоровьем

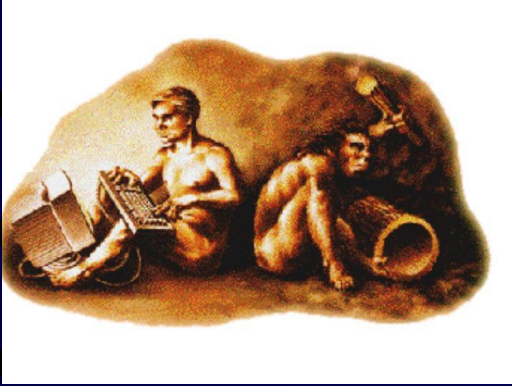




Контроль информационной перегрузки

- Рекомендации по предупреждению *интернет-аддикции* (К. Янг, Д. Гринфилд)
- Навыки “*многозадачности*”
- Направление информации только тем, кому она необходима
- Суммирование поступающей информации с пропуском некоторых деталей, которые можно запросить отдельно (дайджесты, обзоры, подборки)
- Сортировка и выборочная рассылка информации
- Исключение негативной избыточной информации

Социальные, культурные, политические и географические барьеры



- *Цивилизационный кризис Тоффлера (1978)* – увеличение разрыва между содержанием базовых знаний и быстро изменяющимися потребностями общества в новых знаниях, умениях и навыках
- *Системный кризис развития Кумбса (1968)* – проявления глобального кризиса геополитического уровня, состоящего в проблеме обеспечения равных возможностей в получении образования.
- Цифровое неравенство (*digital divide*) в отношении обеспечения доступа к ИКТ между различными группами населения в странах мирового сообщества

Наполнение онлайн-баз

- **1960**, США - собрания э-документов в военных, академических и образовательных организациях, промышленных корпорациях (формат ASCII-файлы без элементов форматирования, *правовая и экономическая области*)
- **1990** - э-тексты с иллюстрациями, графиками и таблицами, идентичные бумажным аналогам (форматы TeX, PostScript, PDF, *естественные, точные и гуманитарные дисциплины*)
- **2000** - редкие и старопечатные книги, полные собрания сочинений античных и средневековых авторов, поэтические антологии
 - *Бесплатные* электронные коллекции текстов
 - *Коммерческие* полнотекстовые БД

Открытые электронные коллекции

- **Формируются** энтузиастами или сообществами, заинтересованными в продвижении собственных взглядов
- **Методы комплектования:** копирование с других сайтов, присылка добровольными дарителями, сканирование и распознавание печатных оригиналов, получение непосредственно из издательства или от авторов электронных копий (camera-ready)
- **Особенности:** невысокое качество текста (грамматические ошибки, без идентичности печатному оригиналу), неполнота, случайность представленных изданий, невысокая эффективность систем навигации и поиска
 - *Проект Гутенберга, 1971* (известные художественные произведения, научные трактаты и публицистические произведения на английском языке)
<http://www.gutenberg.net>
 - *Библиотека Мошкова, с ноября 1994* <http://www.lib.ru/>

Коммерческая информация

- Сведения о компаниях, фирмах, организациях, корпорациях (сайты-визитки, информационные сервера или порталы, электронные магазины)
- Сведения о состоянии мировой экономики и экономики отдельных стран (профессиональные платные базы информационно-аналитических агентств мира)
- Сведения о состоянии отраслевых рынков (специализированные маркетинговые и консалтинговые агентства и службы фирм и организаций, многопрофильные и отраслевые журналы, например "Эксперт", "Профиль")
- Деловые новости (зарубежные профессиональные базы LEXIS-NEXIS, Dialog, Reuters, отечественные агентства "Интегрум-Техно", РИА "Новости", "ИТАР-ТАСС", "Интерфакс", "Gazeta.ru")
- Справочная информация (списки веб-сайтов компаний, телефонно-адресный справочник "Желтые страницы" - Yellow Pages, расписание движения транспорта и пр.)

История создания справочно-правовых систем

- Вторая половина 1960-х (CREDOC , Бельгия, 1967)
- LEXIS-NEXIS, 1967, США -- полнотекстовая профессиональная юридическая информация
 - 1980, 1981 – включение британского законодательства, судебных прецедентов
- 1975 -- Научный центр правовой информации, СПС «Эталон», СССР

Ведущие справочно-правовые системы

- Юридическая справочно-информационная система, ЮСИС, Intralex, 1989
- Гарант, 1990 (+ *Гарант-сервис* – распространение СПС) --международные и федеральные док-ты, судебные решения, финансовые консультации, тексты указов, судебная и арбитражная практика, международные договоры, проекты законов, формы док-тов, бизнес-справки и пр. **Еженедельное пополнение 8—12 тыс. док-тов**
- КонсультантПлюс, 1992 -- федеральные и международные док-ты, судебные решения, финансовые консультации, включает федеральную, региональную и местную БД
- Кодекс, 1991 — для бухгалтеров, юристов, кадровиков, специалистов в области медицины и здравоохранения, малого бизнеса, 12 млн док-ов
- РосПравосудие, 2012 -- судебные решения судов общей юрисдикции, мировых и арбитражных судов РФ (55 млн судебных актов)



Электронное правительство *e-Government*

Способ предоставления информации и оказания уже сформировавшегося набора государственных услуг гражданам, бизнесу, другим ветвям государственной власти и государственным чиновникам, при котором личное взаимодействие между государством и заявителем минимизировано и максимально возможно используются ИТ

Этапы создания э-правительства



РЕГИСТРАЦИЯ НА ПОРТАЛЕ ГОСУСЛУГИ



С помощью портала Госуслуги можно:



Оплачивать коммунальные услуги и штрафы

Записать ребенка в детский сад



Записаться на прием к врачу



Получить загранпаспорт



Проверить свой счет в Пенсионном фонде

Поставить на учет ТС



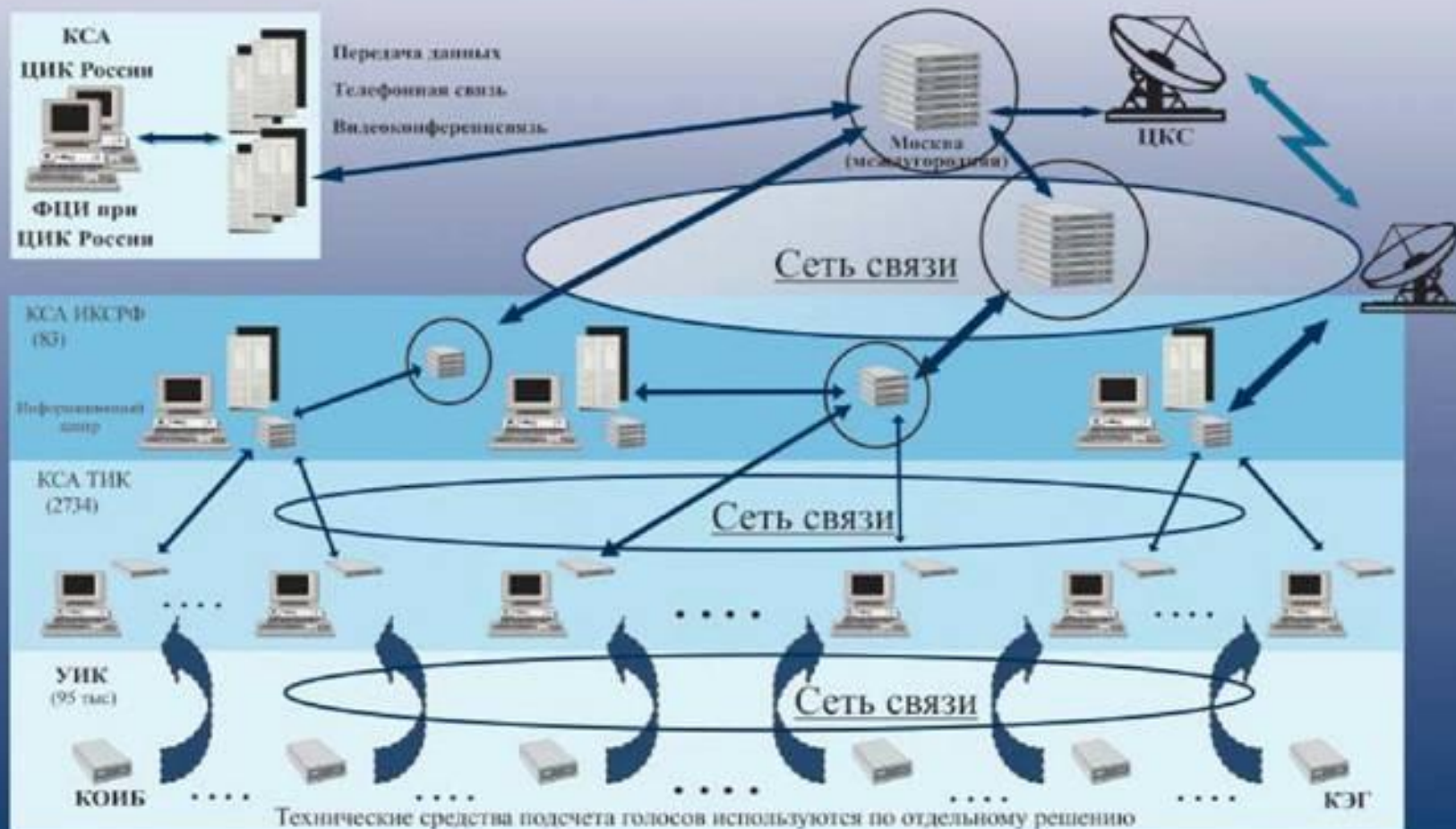
и многое другое.



Автоматизация избирательных действий



Схема организации функционирования ГАС «Выборы»



Вопросы к лекции

1. Что представляет собой рынок информационных услуг? Какие виды информации он включает?
2. Каким образом и благодаря чему увеличивается объем информации? Какие технологические изменения происходят в Интернете? Что мешает их развитию?
3. Какие источники информации популярны сегодня? В чем отличие электронного журнала от печатного?
4. В чем состоят ограничения в переработке информации человеком? Что обозначает интернет-аддикция?
5. Назовите этапы в создании и накоплении данных в открытых и коммерческих электронных (полнотекстовых) коллекциях (БД)? Приведите примеры ведущих справочно-правовых систем