



КРЫМ 2012
МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ

**РЕАЛИИ
ПРАКТИЧЕСКОЙ
ОЦИФРОВКИ**

Специальная
программа

***Комплексный
подход
к
качественной
оцифровке
фондов***

Тимиргалиев Сергей
Махмутович -
Генеральный директор
ЗАО «ДиМи-Центр»

**Международная научно-
практическая конференция**

**«Оценка и контроль
качества сканирования
бумажных документов»**

**г. Королев, Московская
область, 19-20 апреля 2011
года**



Проблемный семинар и круглый стол

**«Нормативно-
технические и
технологические
аспекты создания
электронных копий
документов»**

**г. Судак, Крым,
9 июня 2011 года**



«Необходимо констатировать факт, что на сегодня **не существует** нормативов и методик, которые позволили бы с достаточной степенью уверенности сформулировать техническое задание на оцифровку и организовать сам процесс»

Международная научно-практическая конференция
г. Королев, 2011 год

«У нас в России и во многих других странах **отсутствуют** национальные, регламентирующие показатели качества оцифровки документов»

Шрайберг Я.Л. из ежегодного доклада на Конференции «Крым» год 2011



Сформированы рабочая группы перед которой были поставлены следующие задачи:

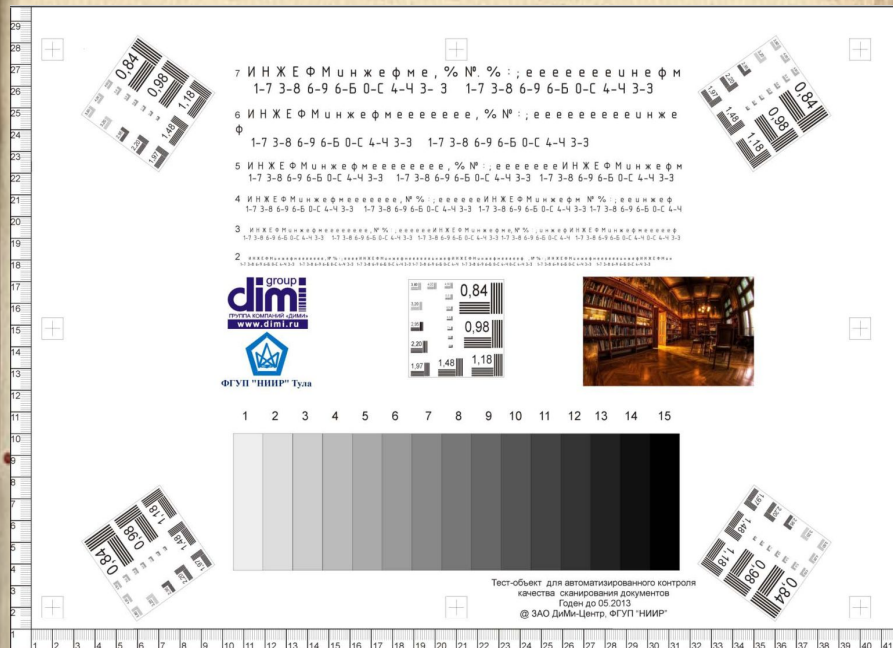
- Анализ мирового опыта;
- Выбор методов контроля;
- Определение категорий качества;
- Определение показателей качества;
- Разработка инструментальных средств контроля качества.

Результаты проданной работы за год:

Универсальный тест-

объект для
автоматизированного
контроля качества
цифровых копий
документов в процессе
сканирования.

Программа
автоматизированного
контроля процессов
сканирования.



Ярк.	D	D*	dD*	CKO
212,46	0,05	0,05	0,038	1,56
194,45	0,15	0,09	0,034	1,94
179,90	0,25	0,12	0,027	2,15
168,90	0,35	0,15	0,022	2,28
160,45	0,45	0,17	0,042	2,34
145,71	0,55	0,21	0,038	2,46
133,47	0,65	0,25	0,038	2,38
122,18	0,75	0,29	0,048	2,59
109,49	0,85	0,34	0,050	2,73
97,59	0,95	0,39	0,048	2,94
87,29	1,05	0,44	0,080	2,95
72,53	1,15	0,52	0,069	2,85
61,83	1,25	0,59	0,121	3,28
46,78	1,35	0,71	0,087	3,12
38,27	1,45	0,79		2,84

Яркость

Разрешение (dpi): 399,9

Кол-во плашек: 1 x 15

Закрашивать

Шкала серого

Расчётная ширина диапазона оптических плотностей: 0,66

Средняя ширина "ступеньки": 0,050

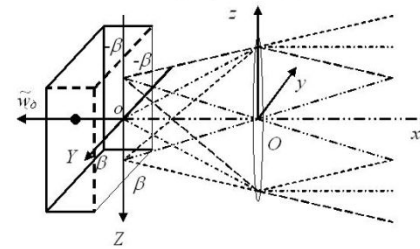
Неравномерность, %:

Результаты проделанной работы за год:

Математическая модель
аналогово-цифрового
преобразования.
Оцифровка документов на
бумажных носителях.

Методика контроля
качества цифровых копий
документов, полученных
их сканированием.

6. Нормированный импульсный отклик фоточувствительной ячейки фотоэлектронного преобразователя



$$\tilde{w}_0(Y, Z) = \begin{cases} \kappa_0 & \text{при } |Y| \leq \beta, |Z| \leq \beta, \\ 0 & \text{во всех остальных случаях.} \end{cases}$$

$2\beta \times 2\beta$ - размер фоточувствительной ячейки

7. Суммарный нормированный импульсный отклик оптикоэлектронки сканера

$$\tilde{w}_n(Y, Z) = \tilde{w}_0(Y, Z) * \tilde{w}_a(Y, Z) = \frac{1}{\sqrt{2\pi[(0,33d_{\text{сп}})^2 + (0,85\beta)^2]}} \exp\left(-\frac{(Y, Z)^2}{2[(0,33d_{\text{сп}})^2 + (0,85\beta)^2]}\right) = \frac{1}{\sqrt{2\pi b_n}} \exp\left(-\frac{(Y, Z)^2}{2b_n^2}\right)$$



Основные концепции Методики обсуждены на Круглом столе

«Создание качественных
электронных ресурсов в
библиотеках, архивах,
музеях и других
государственных
организациях»

г. Москва, 21 февраля 2012
года ГПНТБ России



Испытания Методики
прошли на базе:

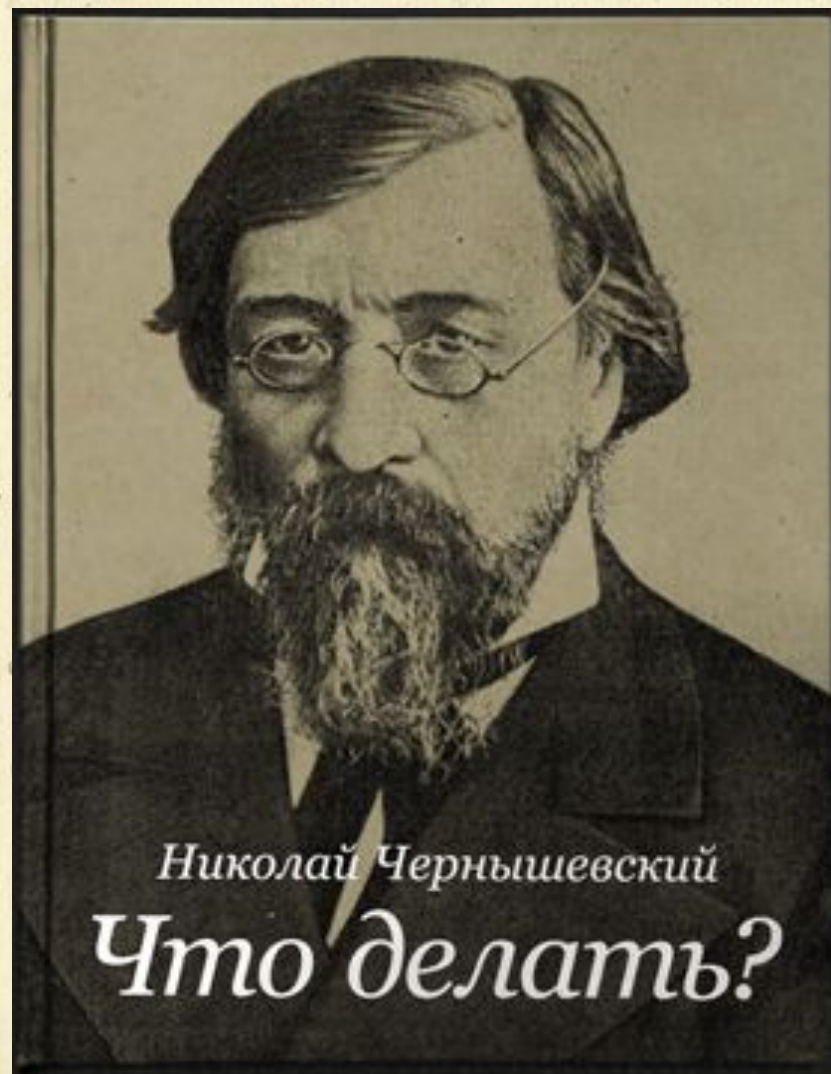
- Государственной
публичной научно-
технической
библиотеки России
- Российской
государственной
библиотеки
- НИИ репрографии
- Тульского
государственного
университета

На основании Методики
разработана первая
редакция

• ГОСТ РО «ЕР СФД.
«Электронные копии
бумажных документов.
Показатели качества и
методы контроля»

**ГОСТ вступит в силу
не ранее 2014 года**

**возникает
естественный
вопрос**



Наше предложение:

Просить Ученые советы:

- НИИ репрографии
- Государственную публичную научно-техническую библиотеку России

- Российскую государственную библиотеку

Одобрить, **указанную**
Методику **и**
рекомендовать **ее**
использовать **до**
появления ГОСТА



Инв. №	Для служебного пользования
Экз. №	
ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ	
НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ	
ОГРАНИЧЕННОГО РАСПРОСТРАНЕНИЯ	
ГОСТ Р 0033 (проект, первая редакция)	
Единый российский страховой фонд документации ЭЛЕКТРОННЫЕ КОПИИ БУМАЖНЫХ ДОКУМЕНТОВ	
Показатели качества и методы контроля	
Настоящий проект стандарта не подлежит применению до его утверждения	
Москва	

**Комплексный
подход
к качественной
оцифровке
фондов**

**Время
бесконтрольного
сканирования
закончилось!**

iPad

12:49 PM

84%



fring



HistoryMaps



Planets



TWC MAX+



Twilight Lite

Наступает эпоха осознанной высокотехнологической деятельности

USA TODAY

Kindle

Marvel

NPR

BBC News



Wikipanion



iPod



Safari



Mail



Photos



TweetDeck



Facebook

iPad

12:49 PM

84%

Мы не можем позволить себе сканировать документ или книгу, имеющие важное историческое или культурное наследие более одного раза.



Wikipanion



iPod



Safari



Mail



Photos



TweetDeck



Facebook

**Необходимо выработать единые
правила оцифровки.**

**Необходим системный подход к
решению комплекса, составляющих
жизненный цикл оцифрованного
документа, включая выбор
оборудования, процесс
сканирования, обработку и хранение
цифровых копий**



iPod



Safari



Mail



Photos



TweetDeck



Facebook

Картина мира оцифровки



Ключевые вопросы оцифровки, на которые в настоящее время

НЕТ однозначного ответа

Сканирование

- Выбор оборудования
- оборудование помещений, освещение
- подготовка материалов;
- параметры сканирования;
- форматы сохранения;
- квалификация сотрудников,
- регламенты работы и контроля сканирования, сопроводительная документация;
- влияние вредных факторов

Обработке исходных сканов

- правила обрезки краев, исправления искажений, корректировки геометрии;
- разрезка разворотов, фрагментация;
 - цифровая реставрация;
 - размерность и масштаб изображений;
 - параметры преобразований форматов;
 - подготовка к распознаванию и распознавание;
 - организация производственных каталогов сканов;

Создания технологических метаданных:

- перечни обязательных сведений о цифровой копии;
- наименования файлов;
- данные про связь с оригиналом и другими копиями;
- правила форматирования технических метаданных;

Формирования полных цифровых копий оригинала:

- типовые наборы копий для каждого вида оригинала;
- форматы файлов;
- комплекты изображений всех представлений оригинала и метаданных на каждую копию;

Создания различных цифровых аналогов документа для разных целей

- функционал и сервисы по контенту;

Вопросы долговременного хранения и оперативного доступа –

- спорными остаются, как правило, дорогостоящие предложения по долговременным носителям.

Многоаспектное использование цифровых аналогов и организация сервисов по контенту

- развитие пользовательского ПО с учетом полно- «визуальных» и полнотекстовых представлений оригиналов и многоязычных полнотекстовых сервисов по контенту.
- что делать с уже оцифрованным материалом плохого качества?
- – пересканирование, переработка?

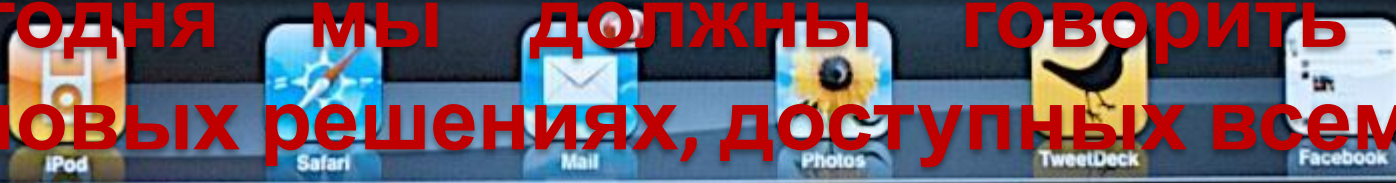
Системы управления на каждом этапе и всего комплекса задач для Оцифрованного документа –

- технико-технологическая задача объективного автоматизированного контроля качества результатов, начиная с организации сканирования и первых технологических процедур оцифровки.

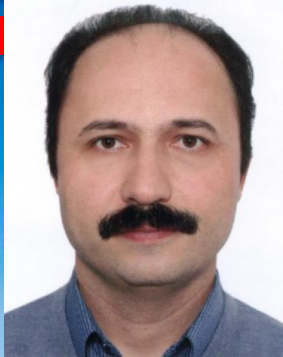
- Примеры таких решений уже есть
- Президентская библиотека им. Б.Н. Ельцина
 - Библиотека конгресса США
 - Фирма «MicroFormat», Нидерланды
 - Корпорация «Элар», Россия и д.р.

Но эти решения являются либо националь-ными проектами, либо очень дорогие, либо закрытые, либо «know how».

Сегодня мы должны говорить о типовых решениях, доступных всем!



Инициативная группа готова к решению НОВЫХ задач и приглашает всех к



iPod



Safari



Mail



Photos



TweetDeck



Facebook

Спасибо за внимание!

Тимиргалиев С.М.

Тел. (495) 780-44-33, (495) 780-44-66

E-mail: tcm@dimi.ru



iPod



Safari



Mail



Photos



TweetDeck



Facebook