



ancestry™

Организация хранения и использования
электронных документов

Николай Доницки – Контент Директор, Европа

Ноябрь 2019

ancestry.com™

8 сайтов & 10 подразделений

 ancestry.co.uk™

 ancestry.ca™

 ancestry.com.au™

 ancestry.de™

 ancestry.it™

 ancestry.se™

 ancestry.fr™

 ancestry.mx™

 ancestry.com™
library edition

 ancestryDNA™

Archives™

fold3 

myfamily.com™

Newspapers.com


rootsweb

FIND A GRAVE

 ProGenealogists™
Official Ancestry.com research firm

 ancestryAcademy™

 ancestry

О КОМПАНИИ

30 +

international
markets

3

million
paying
subscribers

~ 10

million
people DNA-
tested

ЗАПИСИ И КОНТЕНТ

20

billion
records

80

countries
of origin

100

million
family trees

11

billion
connections

ПРОДУКТ И ТЕХНОЛОГИИ

~10

million
people DNA-
tested

75

million
searches
daily

16

million
app
downloads

2017

Webby
Award

НАША КОМАНДА

1600

employees

7

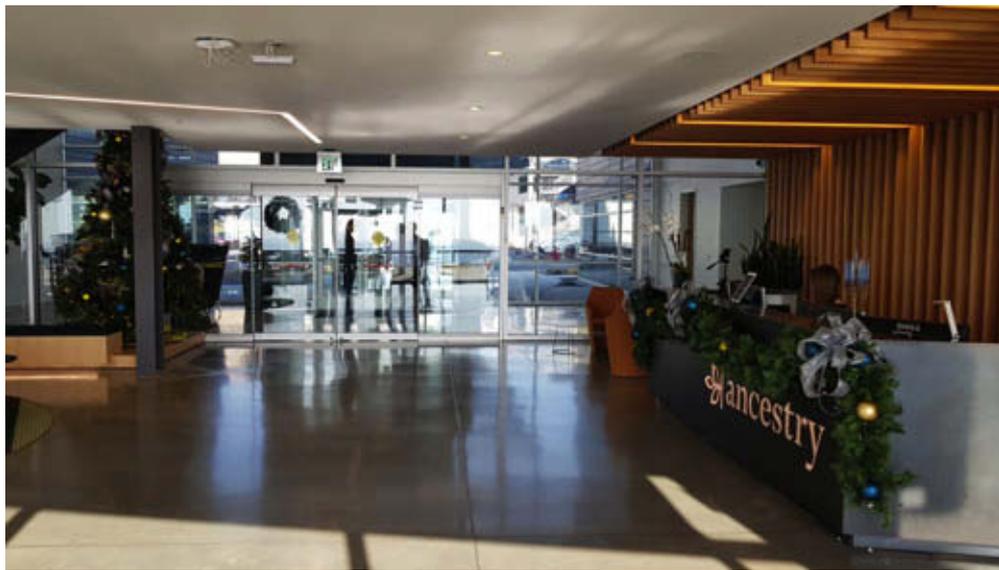
offices
worldwide

Best

Place
to Work
award

Страны в которых мы активно сотрудничаем с Архивами





Предыстория цифровой обработки документов

- Ancestry.com занимается оцифровкой, размещением и публикацией исторических записей с 1996 года.
- В Ancestry размещено более **10 Петабайт** доступных для клиентов изображений и записей – для этого требуется более 5000 серверов

= 10.000 Терабайт



0.002 TB x 5.000.000



= 5,6 x

около 8.900 метров



Некоторые из наших Архивов - Партнеров



TRIBUNALE DI COMO

TRIBUNALE di VERBANIA

TRIBUNALE di CASALE M.TO

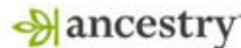
TRIBUNALE DI PAVIA

TRIBUNALE DI LODI



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Schweizerisches Bundesarchiv BAR



Зачем мы храним данные?

Ancestry это крупнейший мировой генеалогический Архив

для ИСТОРИИ. для НАСЛЕДИЯ. для СОХРАНЕНИЯ.

Делиться данными с другими пользователями

Сделать данные доступными и с возможностью поиска



Сбор

Сканирование

Индексация

Публикация

Как Ancestry хранит собственные данные

- В настоящее время все что есть у Ancestry хранится AWS (Amazon Web Services или Сетевой сервис Амазон) Речь идет о безопасном облачном хранении.
- Хранение данных Ancestry соответствует Общему и Генеральному регламенту о защите персональных данных GDPR (European General Data Protection Regulation)
- Защищенность данных также соответствует Политике Privacy Shield, которая строго регламентирует сбор, использование и разглашение персональных данных между США и Странами Евросоюза.



Как мы организуем хранение

Мы следуем международным стандартам, не изобретая собственных:

- Стандартам Немецкого научно-исследовательского общества DFG (German Science Society)
- Стандартам Национальных Архивов США и Великобритании касающимся Наименований папок и файлов.
 - JPEG200 или TIFF, цветной, разрешением 300 dpi
- Стандартам Библиотеки Конгресса США (Library of Congress)
 - METS / Стандарт кодирования и передачи Метаданных
 - MDC / Метаданные для цифрового контента
 - MIX / Метаданные для изображений в XML-схеме
 - AUGIAS, SCOPE, FAUST, STARTEX – CSV (значения разделенные запятой) стандарты индексации базы данных

Сканирование: Источники Оцифровки

- Данные поступают в различном виде:
 - Микрофильмы (16mm и 35mm)
 - Микрофиши
 - Кинги в переплете
 - Отдельные листы
 - В цифровом формате

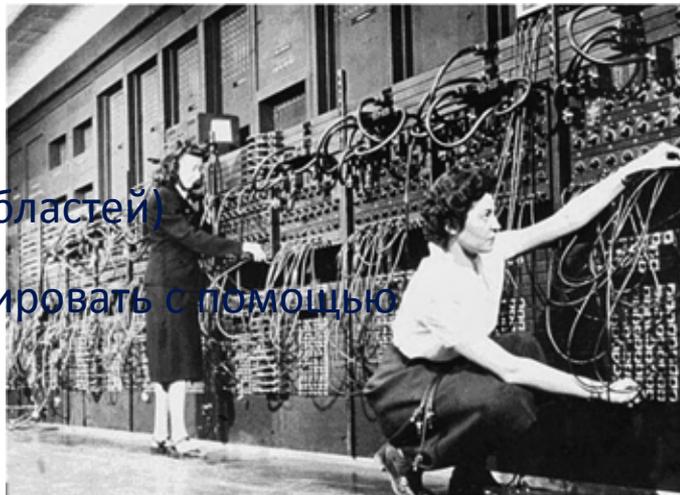


Обработка фотографий

Автоматизированные процедуры уменьшают количество ручных правок :

- Авто - Выравнивание (яркость / контраст)
- Авто – Горизонтальное и вертикальное выравнивание
- Авто - Обрезка (удаление нежелательных областей)

Эти процессы можно автоматически контролировать с помощью инструментов качества Ancestry.



Индексация

Сбор

Сканирование

Индексация

UK Incoming Passenger Lists, 1878-1960 about Cyrus Gold

Name:	Cyrus Gold	 View original image
Birth Date:	abt 1908	
Age:	21	
Port of Departure:	New York, New York, United States	
Arrival Date:	18 Jun 1929	
Port of Arrival:	Southampton, England	
Ship Name:	Leviathan	
Search Ship Database:	View the 'Leviathan' in the 'Passenger Ships and Images' database	
Shipping Line:	United States Line	
Source Citation: Class: BT26; Piece: 913; Item: 33		

Индексирование - это процесс сбора данных, найденных в изображении, в базу данных.



GREGORY	Edith	Am. Exp. Co. Lon.	"	Teacher	62	
GOLD	Cyrus	Handsworth, Birmh. Eng.	"	Missionary	21	
GARRIOTT	Gretchen	Epson Lane Surry Eng.	"	Student		12
GRAWLEY	Mildred	Regent Pl. Hotel, Lon	"	Musician	37	

Профессиональная Индексация



Инструкции для индексации

48227 302022005616 0001-01093

M.P. **11683** Name **NICOLL, Wemyss Eskine A.**

S.S. **"Queen"** Rating **Chief Officer**

Address ~~22 Head St. off Upper Hill St. Liverpool~~
~~6, Poplar Grove, Elms, Liverpool~~

Name of dependant **Cissie A. Nicoll** **Sarah Elizabeth Murray**
(Unmarried wife.)

Relationship **Widow** Date of { death **3rd July**
accident **1918**

Nature of { illness **Drowned**
injury

Award **30 JUN 1921**

M.M.O. **2nd**
M.O. **7**

(x) (369558) Wt.34023/6494 15000 2-21 W & S Ltd.

Service Number

Surname
Given Name

Rank

Ship
Address- *The street should not be keyed*

Relative 2 Given Name
Relative 2 Surname

Relative 1 Given Name
Relative 1 Surname

Death Day
Death Month
Death Year

When it is unclear if it is a death date or an injury date, look for key words such as "Drowned" that also indicate death.

Relation to Head 1

Мастер базы и Словари

- Ancestry создает Мастер базы данных, которые постоянно обновляются и совершенствуются и используются для создания трех типов словарей
 - Ключевые Словари
 - Содержат значения которые могут быть использованы для помощи при вводе.
 - Словари Нормализующие
 - Преобразуют сокращенные обозначения, орфографические ошибки и т. д. в стандартные значения
 - Словари Валидации
 - Содержат стандартные значения

```
<Rel><A>Huntingdon Co.</A><B>Huntingdon</B></Rel>↓  
<Rel><A>Indiana Co.</A><B>Indiana</B></Rel>↓  
<Rel><A>Jefferson Co.</A><B>Jefferson</B></Rel>↓  
<Rel><A>Juniata Co.</A><B>Juniata</B></Rel>↓  
<Rel><A>Lackawanna Co.</A><B>Lackawanna</B></Rel>↓  
<Rel><A>Lancaster Co.</A><B>Lancaster</B></Rel>↓  
<Rel><A>Lawrence Co.</A><B>Lawrence</B></Rel>↓  
<Rel><A>Lebanon Co.</A><B>Lebanon</B></Rel>↓  
<Rel><A>Libanon Co.</A><B>Lebanon</B></Rel>↓  
<Rel><A>Lehigh Co.</A><B>Lehigh</B></Rel>↓  
<Rel><A>Luzerne Co.</A><B>Luzerne</B></Rel>↓  
<Rel><A>Lycoming Co.</A><B>Lycoming</B></Rel>↓  
<Rel><A>Mc Kean Co.</A><B>McKean</B></Rel>↓  
<Rel><A>McKean Co.</A><B>McKean</B></Rel>↓  
<Rel><A>Mercer Co.</A><B>Mercer</B></Rel>↓  
<Rel><A>Mifflin Co.</A><B>Mifflin</B></Rel>↓  
<Rel><A>Monroe Co.</A><B>Monroe</B></Rel>↓  
<Rel><A>Montgomery Co.</A><B>Montgomery</B></Rel>↓  
<Rel><A>Montour Co.</A><B>Montour</B></Rel>↓  
<Rel><A>Northampton Co.</A><B>Northampton</B></Rel>↓  
<Rel><A>Northumberland Co.</A><B>Northumberland</B></Rel>↓  
<Rel><A>Perry Co.</A><B>Perry</B></Rel>↓  
<Rel><A>Philadelphia Co.</A><B>Philadelphia</B></Rel>↓  
<Rel><A>Pike Co.</A><B>Pike</B></Rel>↓  
<Rel><A>Potter Co.</A><B>Potter</B></Rel>↓  
<Rel><A>Schuylkill Co.</A><B>Schuylkill</B></Rel>↓  
<Rel><A>Snyder Co.</A><B>Snyder</B></Rel>↓  
<Rel><A>Somerset Co.</A><B>Somerset</B></Rel>↓  
<Rel><A>Sullivan Co.</A><B>Sullivan</B></Rel>↓
```

Текстовый вид Нормализующего словаря

Публикация: Нормализация

- Нормализация позволяет скорректировать орфографические ошибки и вариации написания в значениях полей.
- Поля которые могут быть включены в процесс нормализации:
 - Имена, Даты, Места, Взаимоотношения, Пол, и многое другое
- Нормализация специфична для разных языков

Примеры Нормализации

- Пол для Английского языка
 - M преобразуется в Male (Мужской), F в Female (Женский)
- Пол для Испанского языка
 - M меняется на Mujer (Женский) и V меняется на Varon (Мужской)
- Возраст для Английского языка
 - 8 mo (8 месяцев) преобразуется в 8/12
- Имя для Испанского языка
 - MA автоматически меняется на María
- Наименование места
 - Utha (Штат Юта, написано с ошибкой) to Utah (верное написание)
- Взаимоотношения для Английского языка
 - Daghter (Дочь, неверное написание) автоматически меняется на Daughter(верно)

Дополнение Данных

- Если данные отсутствуют или недоступны для ввода, они могут быть дополнены
- Дополнение может быть выполнено **помощью:**
 - Предоставленных Метаданных
 - Вычислений
 - Предположений исходя из сопутствующих данных
- Эта дополнительная информация необходима для того, чтобы записи правильно функционировали на веб-сайте.

Примеры Дополнений

- Год Рождения
 - Некто умер в 1850 году в возрасте 45 лет. Дополненные данные: Ориентировочная дата рождения "1805"
- Фамилия
 - Ребенок по имени "Peter" рожденный в штате Массачусетс от отца по имени "John Smith". Дополненные данные: Фамилия "Smith". Имя ребенка "Peter Smith"

Настройка интерфейса пользователя сайта, часть 2

Страница с данными

Панель индексов

All Results SAVE PRINT SHARE

John Francis Smith

in the Northern Territory, Australia, Probate Index, 1911-1994

No image
Text-only collection

Add alternate information Report issue

Name: John Francis Smith
Death Age: 60
Birth Date: abt 1904
Residence Place: Darwin NT/b USA
Death Date: 31 Aug 1964
Death Place: Darwin, Northern Territory
Occupation: Pensioner

SAVE Cancel

Write a comment.

Make a Connection
[Find others](#) who are researching John Francis Smith in Public Member Trees

Source Citation
Genealogical Society of the Northern Territory Inc; Darwin City, Northern Territory, Australia; Northern Territory Probate Index, 1911 - 1994; Reference Number: 2560; File Name: E104 1959-65 Singles; Fobn: E104

Source Information
Ancestry.com. Northern Territory, Australia, Probate Index, 1911-1994 [database on-line]. Provo, UT, USA: Ancestry.com Operations, Inc., 2016.

Original data: Probate Index. Genealogical Society of the Northern Territory Inc.

Description
This collection includes images of an index to probate records for the Northern Territory of Australia. [Learn more.](#)

Name	Death Age	Birth Date	Birth Place	Residence Place	Probate Date	Death Date	Death Place	Occupation	Probate Number	Adjunctum
William John Smith					8 Jun 1921				344	N
Joseph Smith					31 May 1923				407	N
Owen Smyth					7 May 1947				13/47	N
Thomas Smyth					7 May 1947				13/47	N
Owen Smyth					7 May 1947				13/47	N
Charlie Snape					13 Mar 1948				7/1948	N
Samuel Andrew Snape					25 Oct 1953				82/53	N
Lewin Somerville					10 Jun 1936				25/36	N
Al Tough Son					28 Jun 1921				340	N
Wilfred Spicer					9 Jul 1935				39/35	N
Adolph Springer					12 Dec 1912				306	N
Frederick William Spurl					22 Oct 1956				146/56	N
Clarence Reginald Standley					4 Jan 1938				297	N
Arthur William Stanley					25 Jul 1934				18/34	N
Mary Starr			12 mile Darwin, Northern Territory		9 Jul 1913			Married Woman	5/13	N
Grace Isabel Stasinowsky					25 Sep 1952				58/1951	N
Carl Elmer F Steinhilf					21 May 1943				10/43	N
John Stephens				Borroloola, Northern Territory	8 Apr 1923			Stockman	87	N
Alexander Keimann Stevens				Darwin, Northern Territory	7 Aug 1916			Laborer		N
Florent Stevens					16 Aug 1945				18/1945	N
Allan Stevenson					17 Aug 1943				9/43	N
Arthur William Stew					17 Nov 1944				26/44	N
James Pyle Stewart					11 May 1939				16/39	N
Robert Stewart					7 Dec 1948				36/1948	N
Stones					16 Aug 1945				17/1945	N
Stones					15 Jun 1945				13/1945	N

THANKS FOR YOU!

