

POZNAŃ SUPERCOMPUTING AND NETWORKING CENTER



„Introduction to digitisation”

Adam Dudczak

Poznań Supercomputing and Networking Center

maneo@man.poznan.pl



Education and Culture DG

Culture Programme



“Digitisation is the conversion of analogue materials into a digital format for use by software, and decisions made at the time of digitisation have a fundamental impact on the manageability, accessibility and viability of the resources created.”

MINERVA Technical Guidelines for
Digital Cultural Content Creation Programmes

Agenda

- Why digitise?
 - Examples of digitisation projects
- Introduction to digital images
 - Digital image anatomy
 - File formats
 - Software used in graphical post-processing
- Digital text
 - Optical Character Recognition
 - File formats

Agenda (2)

- A few words about digitisation hardware and software
 - Including a quick tour through the e-learning course
- Case study: „OCR in Poland”
- Discussion

Why digitise?



Digitisation of Codex Gigas, picture by Jens Östman (National Library of Sweden)
<http://www.flickr.com/photos/25300312@N08/2960776263/>

Why digitise?



Europeana Remix

An interactive experience around the story of an unlikely friendship during the First World War.



Deutsch

<http://remix.europeana.eu>

Henry III Fine Rolls Project



<http://www.finerollshenry3.org.uk/>

Benefits?

- Increased visibility!
- Rolls were translated into modern English,
- Searching facility was set up to allow users to search for people, places and given subjects,
- Scanned images were released to facilitate a study on the original Latin content,
- Translations with suitable indexes were published as a set of books

Discover Domesday

DOMESDAY BOOK - THE GAME

It is 1086. Last year William the Conqueror ordered a vast survey of land in England to find out who owns the land and how much it is worth to him as King!

There is now a dispute over the numbers of farm animals recorded at this Domesday holding. To give accurate information at a county court you need to count the animals...

COUNT ALL OF THE ANIMALS BY CLICKING THEM - TRY NOT TO MISS ANY



BUT **DON'T COUNT THE FOXES!**



START GAME

<http://www.nationalarchives.gov.uk/domesday/>

Benefits?

- Popularity growth
- Base for interesting teaching material
 - Domesday book – THE GAME
- Better understading of history

News Paper Article
trove.nla.gov.au/ndp/del/article/18151856

Trove Digitised newspapers and more
Home About Browse Help

It's free and it only takes a minute Sign up Login

Advanced search

The Sydney Morning Herald (NSW : 1842-1954) about Tuesday 14 March 1950 Page 1 of 15 Send to

No. 35,016

LABOUR M.L.A. QUILTS PARTY

A.L.P. Split Widens

Mr. F. Stanley, M.L.A., resigned from the Labour Party last night, and, followed by 37 supporters, he walked out of a meeting of the Lakemba branch of the A.L.P.

Later he said he would stand as an Independent Labour candidate at the State elections.

North Sydney A.L.P. branch last night decided to defy the party's central executive and endorse the sitting member, Mr. J. L. Geraghty as its candidate at the elections.

These were the most dramatic developments yesterday in the split in the Labour Party caused by the executive's decision not to endorse four members of the Legislative Assembly.

LAKEMBA MEMBERS RESIGN

The developments last night were the result of the central executive's refusal to endorse four Labour M.L.A.s.

They are: Messrs. Stanley, Geraghty, R. J. Heferen (Barwon) and J. W. Seiffert (Monaro).

On Sunday the Barwon A.L.P. electorate council gave the central executive one week pre-selection ballot to replace me.

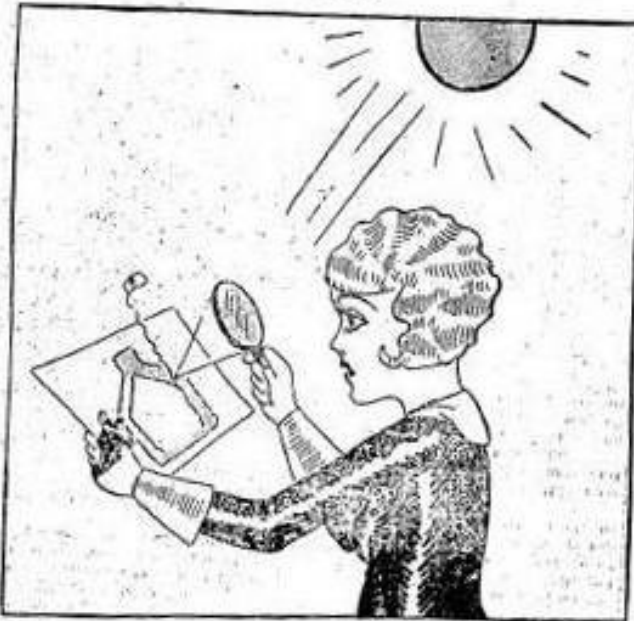
"As this was in direct opposition to the motion that had just been carried, there was a storm of protest.

"About 46 of the 70 members present placed their resignations on Mr. O'Malley-Jones' table in protest.

Scene of the world's worst air disaster, near Llandaff airport, Carr Tudor V, was coming from Dublin carrying Welsh footballers.

ZOOM

How To Draw Pictures: With The Sun As A "Pencil."



Labour M.L.A.s.

They are: Messrs. Stanley, Geraghty, R. J. Heferen (Barwon) and J. W. Seiffert (Monaro).

On Sunday the Barwon A.L.P. electorate council gave the central executive one week

ht
ecutive
J. L.
ions.
p-
bour
ion not
ative

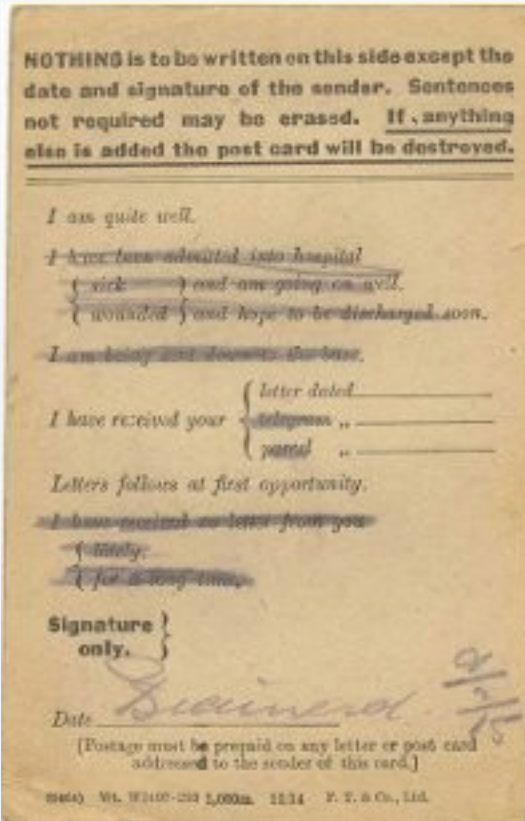
Australian Newspapers Digitisation

<http://trove.nla.gov.au/newspaper>

Benefits?

- Did you know?
 - „Convicts could leave the colonies after their sentences were completed or after being granted an Absolute Pardon. Departures were announced in the Sydney Gazette's 'Notice of Intent' column.”
- In general... increased visibility! ;-)
- Great source of knowledge about genealogy

The Great War Archive



<http://www.oucs.ox.ac.uk/ww1lit/gwa/>

Benefits?

- Mobilization of local community
- Chance to preserve stories and unique memorabilia of particular people
- Great educational material
- Not mentioning... increased visibility

Cultural heritage and technology



Cultural heritage and technology



Cultural heritage and technology

The image displays four screenshots of the Timebook application interface, which is designed to explore historical personalities and their cultural heritage. Each screenshot shows a search bar, a navigation menu, and a main content area.

- Top Left Screenshot:** Shows the 'Timebook' interface with a search bar and a navigation menu. The main content area features a large clock graphic with the numbers 'NINE', '12', and 'THREE' on the hour markers, and '6' on the minute marker. Below the clock are icons for 'Wall', 'Info', 'Photos', 'Music', and 'Friends'.
- Top Right Screenshot:** Shows the 'Timebook' interface for 'William Blake - Photos'. The main content area displays a grid of images, including a handwritten manuscript and a painting of a landscape. A navigation menu is visible on the right side.
- Bottom Left Screenshot:** Shows the 'Timebook' interface for 'Edouard Manet'. The main content area displays a grid of images, including a portrait of Manet and a painting of a woman. A navigation menu is visible on the left side.
- Bottom Right Screenshot:** Shows the 'Timebook' interface for 'Pablo Picasso - Photos - Works in Europeana'. The main content area displays a grid of images, including a portrait of Picasso and a painting of a woman. A navigation menu is visible on the right side.

The interface also includes a 'Time machine' feature, as seen in the bottom right screenshot, which allows users to explore the profiles of personalities of the past, explore their interest and comment their statuses. A specific example shown is 'Beethoven' with the text 'See his profile' and a portrait of the composer.

Summary of benefits

- Provides better access to unknown or little used collections,
- Offers better search and retrieval facilities for an image collection,
- Provides a better understanding of original works through improved indexing or some form of digital image enhancement,
- Creates resources that are suitable for use in learning and teaching,
- Enhances the public knowledge, recognition and understanding of cultural heritage,
- Facilitates preservation of collections for the future in a digital format.

Agenda

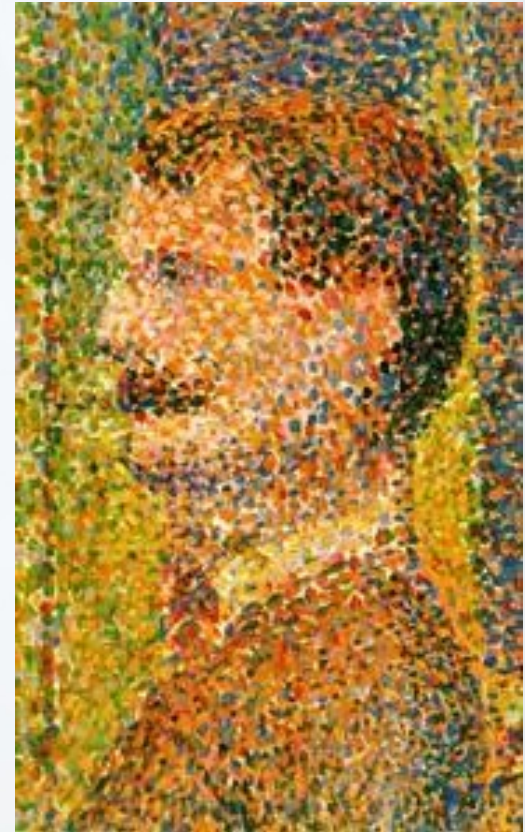
- Why digitise?
 - Examples of digitisation projects
- **Introduction to digital images**
 - Digital image anatomy
 - File formats
 - Software used in graphical post-processing
- Digital text
 - Optical Character Recognition
 - File formats

Digital image anatomy

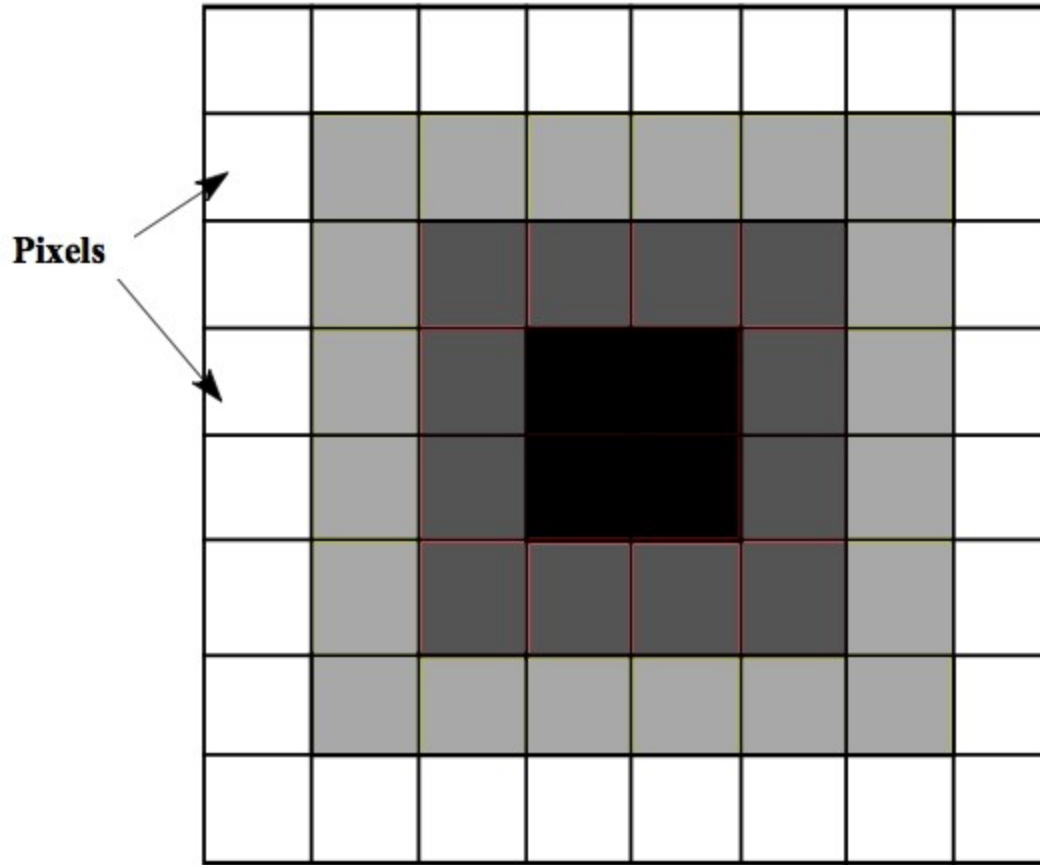
„Digital image is a numeric representation (built out of 0s and 1s) of a two-dimensional image”

Wikipedia

- Raster images
- Pixels?



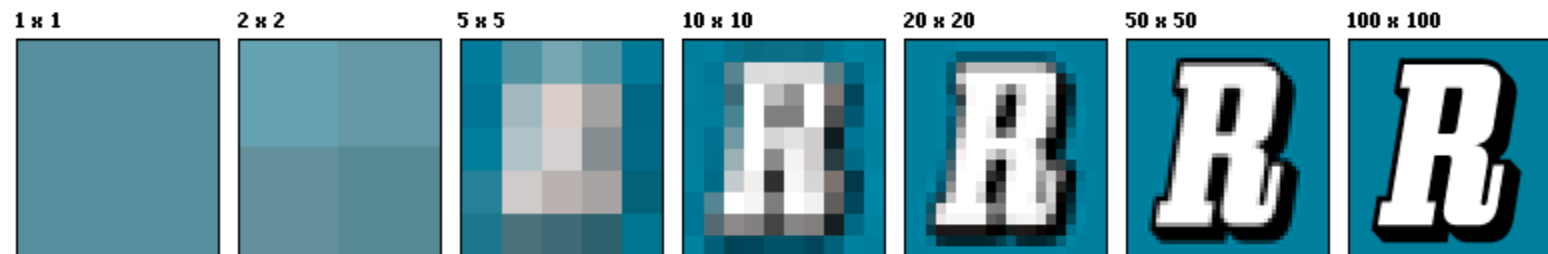
Digital image anatomy (2)



- Each pixel:
 - holds information about color
 - has certain position in the image
- Higher number of pixels means better quality of image

Image resolution

- Resolution is a term that describes image ability to hold information
- Higher resolution means that image will be more detailed
- Resolution can be expressed in three ways



Number of pixels

- Width and height of an image expressed in pixels
 - e.g. 640 x 480
- Sum of all pixels expressed in megapixels
 - e.g. 0.3 megapixel means that image consist of 307200 pixels, it is the same $640 * 480$ image

Color depth

- Each pixel holds information about its color
 - This information can be less or more detailed
- Color depth denotes accuracy of color representation
- It is expressed in bits per pixel (Bits per pixel, BPP)

Color depth (2)

- 1 BPP – only two possible colors
- 24 BPP – 16,7 milion of colors

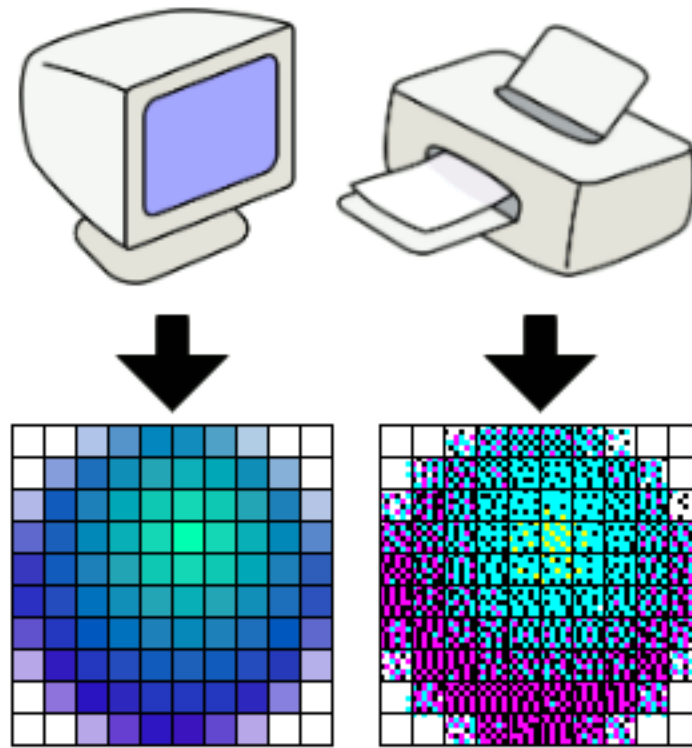


PPI

- Image resolution can be expressed using pixel density per inch of [screen, print-out]
- PPI = Pixel Per Inch
- Higher number means more detailed scan/print out

DPI vs PPI

- Dots and pixels are terms from two different worlds



DPI vs PPI (2)

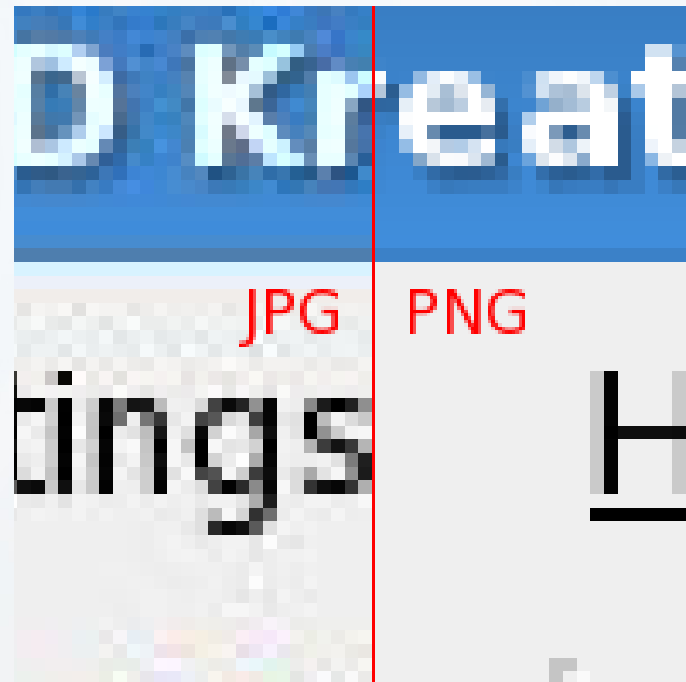
- Digital image quality should be expressed using PPI
- During scanning you need to use Dot Per Inch (?!?)
- In practice this units of measure are used interchangeably

Image compression

- We would like to have very detailed representation of original image, but:
 - we don't like long download times,
 - we don't have money to buy numerous DVDs, required to archive our digital copies
- Image compression can reduce size of graphical files

Image compression (2)

- Lossless compression
- Lossy compression



Lossy compression

- It is a compromise between information loss and file size reduction
- Usually there is a way to adjust compression ratio
 - Higher compression ratio means that more information is lost but the file is smaller
- Lossy compression is irreversible process
 - Lost information can't be recovered even after decompression
 - This is crucial while working with images!!
- This way of compression is usually more efficient than lossless compression
- Examples of pictures compressed using lossy compression:
 - <http://www.bbproject.net/2010/05/jak-wyglada-jpeg-po-100-000-rekompresji/>

File formats – what is important?

- Does it impose compression type/method?
- Does it allow to embed metadata in an image?
- Can one file consist of multiple images?
- Is specification of file format publically available?
- During this training we will focus on three most widely used gfx file formats: TIFF, JPG, PNG

TIFF

- Tagged Image File Format
- Available since 1986
- Specification is available for everyone
- It allows to:
 - Embed metadata,
 - One file can contain multiple images
 - It can be used without any compression as well as with lossy and lossless compression
- TIFF (+ lossless LZW compression) is considered as an de facto standard in digital archiving

RAW

- RAW is a file format containing unprocessed data from camera image sensors
- Each vendor has its own, so it requires dedicated software
- It allows for advanced adjustments of an image
- Sometimes called digital negative

JPEG

- **J**oint **P**hotography **E**xperts **G**roup
- File format created especially for storing photography
- Very widely used
- It uses lossy compression!

PNG

- **P**ortable **N**etwork **G**raphics
- General purpose file format was suppose to replace GIF files
- It employs lossless compression
- PNGs are usually bigger than JPEGs

Interactive formats

- Allows to dynamically adjust the quality of presented image
- Especially useful for large images which had to be delivered in high level of details e.g. maps
- Three major formats Zoomify, Deep Zoom and Google Maps
 - e.g. <http://sbc.wbp.kielce.pl/publication/8855>

Software

- Gimp (<http://gimp.org>)
- IrfanView (<http://www.irfanview.com>)
- ImageMagick (<http://www.imagemagick.org>)
- OptiPNG (<http://optipng.sourceforge.net>)
- ZoomifyExpress (<http://www.zoomify.com>)
- More to come...

Agenda

- Why digitise?
 - Examples of digitisation projects
- Introduction to digital images
 - Digital image anatomy
 - File formats
 - Scanning resolution
 - Software used in graphical post-processing
- **Digital text**
 - Optical Character Recognition
 - File formats

Digital text

What is the difference between scanned text and digital text?



Source: <http://www.flickr.com/photos/shokai/3255725751/>

Why digital text?

- Full text search
 - Even most detailed metadata cannot replace described object
 - Full text search allows to access to full content of the object
 - Crucial for regular and address books
- Reuse
 - Creation of translation, digital editions, monographies is simpler if you can just copy and paste the text
- Research material
 - e.g. Linguistic analysis on how language was changing

Creation of digital text

- Manual keying
 - Manual rewriting of text from scans done by human
 - Expensive (very time-consuming)
 - It may require some domain knowledge
 - Crowdsourcing
 - Trove - <http://trove.nla.gov.au/newspaper>
 - PG DP - <http://www.pgdp.net/c/>
 - Transcribe Bentham - <http://www.ucl.ac.uk/transcribe-bentham/>
 - Wikisource – <http://wikisource.org>
 - Flaws of crowdsourcing

Creation of digital text

- Optical Character Recognition (OCR)
- Computer program converts scanned text into machine-readable digital text
- Quality of results depends on multiple factors

How does OCR works?

- Pre-processing
 - Depending on input image type conversion might be required
 - Binarization, deskewing
- Layout analysis and identification of text components on the page
 - It is not always supported by OCR, crucial for complex layouts like newspapers
- Segmenetation of text into lines, words and characters

FAMILIE SZLACHECKIE.*

L.

Labady. *Indig.* w konse. r. 1700 fol. 227. *Maloch.*
Lach Szyrma h. *Lubisz* z odmiann. *Patrz Szyrma.*
Lachucki w wilkomińskim h. *Maloch.* — Ignacy memnian gub. grodzieński, pułkownik wojsk polskich, wydał *Biogrofa ułoczanina* i przedmoważy tragedya Woltera „Alasty” (Warsz. 1780). Ignacy Emanuel (ur. 1793 i 1825) synowie poprzedniego, dr filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki Czechowicz. *Kuryp.* i *Maloch.* Obacz Czechowicz.
Lachowicki h. Łodziez. w urzaskim pow. Z tych Piotr podpisał elekcyą Jana III. *Niec.*
Lachowicki h. Sienista. Jan syn Aleksandra stolnika wielkiego 1701 i Katarzyny Gerwanickiej, burdyscy krzemieniecki, miał za sobą *Jmo soto Maryanna Rzeczycka*, *Szo soto Maryanne Bąkowska*; z pierwszój był syn Felix, którego z Klary Robożistkiej córka Maryanna poślubiła *Adela Jelwickiego. Gen. Jelwicki.*
Lachowicz h. Łodziez. Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1809, to samo *Wojciech* (als herb jego nie wyrażony) w sędzie z Iwowskiem r. 1783. — Do tego domu zdaje się należeć Stanisław August, współczesny autor w Petersburgu.
Lachowicz h. Radziszew. *Kuryp.* i *Maloch.*
Lachowski h. Nolecz. Tadeusz i Mikolaj udowodnił szlachectwo w sądzie z przemyskiem 1782.
Lachowski h. Pilewa. w sandomierskiem i innych województwach. N. miał za sobą Zofia księżną Waronicką około r. 1520. *Wojciech* wojski z. wiskiej, z sejma 1591 komisarz do granic pruskich ob-Mazowsza. N. pojął Eukowską, córka Jana pisarza z krakowskiego, z Zofii Palcowskiej urodzon. Jan w sandomierskiem 1632. Jednost komornik czerwi 1674. *Niec.* — Niewiadomo czy do tych należał Sebastian Jozua, sławny karczodziela króla Stanisława Augusta. *Kwaz.*
Lacki h. wś. Autorowie nasi piszą, że dawniejszy przodek tego domu Zawichojski nazwany, z Polski wysiedliły, w Moskwa był osiadł, gdzie nazwany był Lackim, a *Dobosier* w kasaniu *Wyprowadzenie* *Obły* przydaje, że ten Zawichojski Zacharyasz podjął z Polski do Moskwy, tam był osiadł; syn jego, wódz z. rycerskich, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który uchwycił zwrócić tyranii Owerpny opiekuna młodego cara Iwana, do Litwy się przeniósł za Zygmunta i króla polskiego, od którego wsiad dobra *Zobudek* i *Wysocki Dwór*, był oraz leśniczym podlaskim, miał wojenny i w sprawach rycerskich biegły, z Bogdanem Polubińskim spłodził córkę Teodorę z Eliaszem Piłgrynowskim pisarzem lit. i synów dwóch: Grzegorza i Teodora, obadwa byli dowódcami kawalerickimi walczeni, obawili jednak Teodor, którego krewne i przeważnie dzieła z pochwałą wspomina *Konary 1611 fol. 45 i 1613*; umarł z wielkim majątkiem i rycerskich w Wilnie, zastawiony z Izabellą Bonarrelli de Rovere, szlachnej wo Włoszech paronitki, córkę Kawarazę z Teodorem Tyszkiewiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę smoleński, radziński, i zosiński, również wojennego meca, ten fundował OO. Dominikanów w Wysockim Dworze; żona jego Jazna Talwoszczona, jedynacka córka Adama karczalana smoleńskiego, z której oprócz córki Izabelli Jmo soto za Teodorem Tyszkiewiczem wojewodą brasławiańskim, *Szo soto* za Hieronimem Chodkiewiczem starostą, mozyrkim, zostało synów trzech: Teodor, Samuel i Piotr;

CKIE.*

podlaski, pułkownik wojsk „Alasty” (Warsz. 1780) i, był redaktorem kilku czasopism w Wilnie.
 701 i Katarzyny Gerwanickiej, burdyscy krzemieniecki, miał za sobą *Jmo soto Maryanna Rzeczycka*, *Szo soto Maryanne Bąkowska*; z pierwszój był syn Felix, którego z Klary Robożistkiej córka Maryanna poślubiła *Adela Jelwickiego. Gen. Jelwicki.*
 1809, to samo *Wojciech* (als herb jego nie wyrażony) w sędzie z Iwowskiem r. 1783. — Do tego domu zdaje się należeć Stanisław August, współczesny autor w Petersburgu.
 1782.
 1520. *Wojciech* wojski z. wiskiej, z sejma 1591 komisarz do granic pruskich ob-Mazowsza. N. pojął Eukowską, córka Jana pisarza z krakowskiego, z Zofii Palcowskiej urodzon. Jan w sandomierskiem 1632. Jednost komornik czerwi 1674. *Niec.* — Niewiadomo czy do tych należał Sebastian Jozua, sławny karczodziela króla Stanisława Augusta. *Kwaz.*
 1611 fol. 45 i 1613; umarł z wielkim majątkiem i rycerskich w Wilnie, zastawiony z Izabellą Bonarrelli de Rovere, szlachnej wo Włoszech paronitki, córkę Kawarazę z Teodorem Tyszkiewiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę smoleński, radziński, i zosiński, również wojennego meca, ten fundował OO. Dominikanów w Wysockim Dworze; żona jego Jazna Talwoszczona, jedynacka córka Adama karczalana smoleńskiego, z której oprócz córki Izabelli Jmo soto za Teodorem Tyszkiewiczem wojewodą brasławiańskim, *Szo soto* za Hieronimem Chodkiewiczem starostą, mozyrkim, zostało synów trzech: Teodor, Samuel i Piotr;

FAMILIE SZLACHECKIE.*

L.

Labady. *Indig.* w konse. r. 1700 fol. 227. *Maloch.*
Lach Szyrma h. *Lubisz* z odmiann. *Patrz Szyrma.*
Lachucki w wilkomińskim h. *Maloch.* — Ignacy memnian gub. grodzieński, pułkownik wojsk polskich, wydał *Biogrofa ułoczanina* i przedmoważy tragedya Woltera „Alasty” (Warsz. 1780). Ignacy Emanuel (ur. 1793 i 1825) synowie poprzedniego, dr filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki Czechowicz. *Kuryp.* i *Maloch.* Obacz Czechowicz.
Lachowicki h. Łodziez. w urzaskim pow. Z tych Piotr podpisał elekcyą Jana III. *Niec.*
Lachowicki h. Sienista. Jan syn Aleksandra stolnika wielkiego 1701 i Katarzyny Gerwanickiej, burdyscy krzemieniecki, miał za sobą *Jmo soto Maryanna Rzeczycka*, *Szo soto Maryanne Bąkowska*; z pierwszój był syn Felix, którego z Klary Robożistkiej córka Maryanna poślubiła *Adela Jelwickiego. Gen. Jelwicki.*
Lachowicz h. Łodziez. Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1809, to samo *Wojciech* (als herb jego nie wyrażony) w sędzie z Iwowskiem r. 1783. — Do tego domu zdaje się należeć Stanisław August, współczesny autor w Petersburgu.
Lachowicz h. Radziszew. *Kuryp.* i *Maloch.*
Lachowski h. Nolecz. Tadeusz i Mikolaj udowodnił szlachectwo w sądzie z przemyskiem 1782.
Lachowski h. Pilewa. w sandomierskiem i innych województwach. N. miał za sobą Zofia księżną Waronicką około r. 1520. *Wojciech* wojski z. wiskiej, z sejma 1591 komisarz do granic pruskich ob-Mazowsza. N. pojął Eukowską, córka Jana pisarza z krakowskiego, z Zofii Palcowskiej urodzon. Jan w sandomierskiem 1632. Jednost komornik czerwi 1674. *Niec.* — Niewiadomo czy do tych należał Sebastian Jozua, sławny karczodziela króla Stanisława Augusta. *Kwaz.*
Lacki h. wś. Autorowie nasi piszą, że dawniejszy przodek tego domu Zawichojski nazwany, z Polski wysiedliły, w Moskwa był osiadł, gdzie nazwany był Lackim, a *Dobosier* w kasaniu *Wyprowadzenie* *Obły* przydaje, że ten Zawichojski Zacharyasz podjął z Polski do Moskwy, tam był osiadł; syn jego, wódz z. rycerskich, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który uchwycił zwrócić tyranii Owerpny opiekuna młodego cara Iwana, do Litwy się przeniósł za Zygmunta i króla polskiego, od którego wsiad dobra *Zobudek* i *Wysocki Dwór*, był oraz leśniczym podlaskim, miał wojenny i w sprawach rycerskich biegły, z Bogdanem Polubińskim spłodził córkę Teodorę z Eliaszem Piłgrynowskim pisarzem lit. i synów dwóch: Grzegorza i Teodora, obadwa byli dowódcami kawalerickimi walczeni, obawili jednak Teodor, którego krewne i przeważnie dzieła z pochwałą wspomina *Konary 1611 fol. 45 i 1613*; umarł z wielkim majątkiem i rycerskich w Wilnie, zastawiony z Izabellą Bonarrelli de Rovere, szlachnej wo Włoszech paronitki, córkę Kawarazę z Teodorem Tyszkiewiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę smoleński, radziński, i zosiński, również wojennego meca, ten fundował OO. Dominikanów w Wysockim Dworze; żona jego Jazna Talwoszczona, jedynacka córka Adama karczalana smoleńskiego, z której oprócz córki Izabelli Jmo soto za Teodorem Tyszkiewiczem wojewodą brasławiańskim, *Szo soto* za Hieronimem Chodkiewiczem starostą, mozyrkim, zostało synów trzech: Teodor, Samuel i Piotr;

*) Herbny rodzin szlacheckich są umieszczone w tomie I.

Deskew

mlaszczona w tomie I

Auto Manual

-1.56°

FAMILIE SZLACHECKIE.

L.

Labady. *Indy* w konst. r. 1790 fol. 227. *Maloch.*
Labę Szyrma h. *Lubiec* z odmian. *Patrz Szyrma.*
Lachnicki, w wilkomirskim h. *Maloch.* — Ignacy ziemianin guberni grodzieńskiej, pułkownik polski, wydał *Biografię waleczności i przelomaczy* tragedya Woltera „Alzjre” (Warsz. 1780). (ur. 1783 + 1820) synowiec poprzedniego, dr. filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki *Czechowicz.* *Kurup.* i *Maloch.* Obraz *Czechowicz.*
Lachowicki h. *Lodzin.* w orsząńskim pow. Z tych Piotr podpisał elekcyą Jana III. *Nie.*
Lachowicki h. *Steniska.* Jan syn Alexandra stolnika wielkiego 1701 i Katarzyny Gorwanienickiej, borsudczy krzemieniecki, miał za sobą *Jano* coto Maryanna Reeczyka, *Żdo* coto Maryanna Bakuwika; z pierwszą był syn Feliks, którego z Klary Rahobinskiej córka Maryanna poślubiła *Gen. Jelowickich.*
Lachowicz h. *Lodzin.* Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1800.
Lachowicz h. *Lodzin.* Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1800. Do tego d. Jan nakaz Stanisław August, współczesny autor w Petersburgu.
Lachowicz h. *Rudawa.* *Kurup.* i *Maloch.*
Lachowski h. *Nolca.* Tadeusz i Mikolaj udowodnił szlachectwo w sądzie z przemyskiem 1782.
Lachowski h. *Pilawa.* w sandomierskim i innych województwach. N. miał za sobą *Zofia* księżna Weroniecka około r. 1520. Wojciech wojak z wiskiej, z sejm 1581 komisarz do granic pruskich od Mazowsza. N. pojął Łukowską, córkę Jana pisarza w krakowskiego, z *Zofia* Palczowskiej urodzoną. Jan w sandomierniem 1632. Jednej komornik czeski 1674. *Nie.* — Niewiadomo czy do tych należał Sebastian Juscia, sławny kasnodzieja króla Stanisława Augusta. *Kroz.*
Lacki h. od Autorowa nasi piza, że dawniejszy przodek tego domu Zawichojki nazwany w przedk. w Moskiewie był osadzi, gdzie nazwany był Lackim, a *Dobosier* w Kazaniu Wyprowadzenie *Gońcy* przysłał, że ten Zawichojki Zacharyasz postąpił z *Polski* do Moskwy, tam był osadzi, syn jego według niektórych, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który niemogąc anielsi tyranii Oweyny opiekuna młodego cara Iwana, do Litwy się przesłał za *Zygnuntia* i króla polskiego, od którego wdał dobra Zoludek i Wysoki Dwór, był oraz leźniczym podkajm, mał wojenny i w sprawach rygarskich biegły, a Bogdana Polubinska spłodził córkę Teodorę za Eliaszem Pielgrzymowskim pisarzem lit. i synów dwóch: Grzegorza i Teodora, obadwa byli działami kawalerskimi wstawiemi, osobliwie jednak Teodor, którego krawce i przeważnie *Gońcy* z pochwałą wspomina *Konasty 1611 fol. 45 i 1613;* umarł z wielkim mająm ludzi rygarskich w Wilnie, zastawiony z Izabelli Bonarulli de Rovere, znaczej we Włoszech paroneli, córkę Katarzynę za Teodorem Pielgrzymowiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę zmadek, raduńsk, i zaleski, również wojennego mecia, ten fundował OO. Dominikanów w Wysekim Dworze; żona jego Joanna Talwosówna, jedynaczka córka Adama kaszelska żmudzkiego, z której opłódz córki Izabelli *Jano* coto za Teodłem Tępną wojewodą brzeskojaskim, *Żdo* coto za Hieronimem Chłkiewiczem starostą, mozyrykim, zostali synów trzech: Teodor, Samuel i Piotr;

Deskw

FAMILIE SZLACHECKIE.*

L.

Labady. *Indy* w konst. r. 1790 fol. 227. *Maloch.*
Labę Szyrma h. *Lubiec* z odmian. *Patrz Szyrma.*
Lachnicki, w wilkomirskim h. *Maloch.* — Ignacy ziemianin guberni grodzieńskiej, pułkownik polski, wydał *Biografię waleczności i przelomaczy* tragedya Woltera „Alzjre” (Warsz. 1780). (ur. 1783 + 1820) synowiec poprzedniego, dr. filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki *Czechowicz.* *Kurup.* i *Maloch.* Obraz *Czechowicz.*
Lachowicki h. *Lodzin.* w orsząńskim pow. Z tych Piotr podpisał elekcyą Jana III. *Nie.*
Lachowicki h. *Steniska.* Jan syn Alexandra stolnika wielkiego 1701 i Katarzyny Gorwanienickiej, borsudczy krzemieniecki, miał za sobą *Jano* coto Maryanna Reeczyka, *Żdo* coto Maryanna Bakuwika; z pierwszą był syn Feliks, którego z Klary Rahobinskiej córka Maryanna poślubiła *Gen. Jelowickich.*
Lachowicz h. *Lodzin.* Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1800. Do tego d. Jan nakaz Stanisław August, współczesny autor w Petersburgu.
Lachowicz h. *Rudawa.* *Kurup.* i *Maloch.*
Lachowski h. *Nolca.* Tadeusz i Mikolaj udowodnił szlachectwo w sądzie z przemyskiem 1782.
Lachowski h. *Pilawa.* w sandomierskim i innych województwach. N. miał za sobą *Zofia* księżna Weroniecka około r. 1520. Wojciech wojak z wiskiej, z sejm 1581 komisarz do granic pruskich od Mazowsza. N. pojął Łukowską, córkę Jana pisarza w krakowskiego, z *Zofia* Palczowskiej urodzoną. Jan w sandomierniem 1632. Jednej komornik czeski 1674. *Nie.* — Niewiadomo czy do tych należał Sebastian Juscia, sławny kasnodzieja króla Stanisława Augusta. *Kroz.*
Lacki h. od Autorowa nasi piza, że dawniejszy przodek tego domu Zawichojki nazwany w przedk. w Moskiewie był osadzi, gdzie nazwany był Lackim, a *Dobosier* w Kazaniu Wyprowadzenie *Gońcy* przysłał, że ten Zawichojki Zacharyasz postąpił z *Polski* do Moskwy, tam był osadzi; syn jego według niektórych, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który niemogąc anielsi tyranii Oweyny opiekuna młodego cara Iwana, do Litwy się przesłał za *Zygnuntia* i króla polskiego, od którego wdał dobra Zoludek i Wysoki Dwór, był oraz leźniczym podkajm, mał wojenny i w sprawach rygarskich biegły, a Bogdana Polubinska spłodził córkę Teodorę za Eliaszem Pielgrzymowskim pisarzem lit. i synów dwóch: Grzegorza i Teodora, obadwa byli działami kawalerskimi wstawiemi, osobliwie jednak Teodor, którego krawce i przeważnie *Gońcy* z pochwałą wspomina *Konasty 1611 fol. 45 i 1613;* umarł z wielkim mająm ludzi rygarskich w Wilnie, zastawiony z Izabelli Bonarulli de Rovere, znaczej we Włoszech paroneli, córkę Katarzynę za Teodorem Pielgrzymowiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę zmadek, raduńsk, i zaleski, również wojennego mecia, ten fundował OO. Dominikanów w Wysekim Dworze; żona jego Joanna Talwosówna, jedynaczka córka Adama kaszelska żmudzkiego, z której opłódz córki Izabelli *Jano* coto za Teodłem Tępną wojewodą brzeskojaskim, *Żdo* coto za Hieronimem Chłkiewiczem starostą, mozyrykim, zostali synów trzech: Teodor, Samuel i Piotr;

*) Herbory rodzin szlacheckich są umieszczone w tomie I.

MILIE SZLACHECKIE.*

L.

konst. r. 1790 fol. 227. *Maloch.*
Labę Szyrma h. *Lubiec* z odmian. *Patrz Szyrma.*
Lachnicki, w wilkomirskim h. *Maloch.* — Ignacy ziemianin guberni grodzieńskiej, pułkownik polski, wydał *Biografię waleczności i przelomaczy* tragedya Woltera „Alzjre” (Warsz. 1780). (ur. 1783 + 1820) synowiec poprzedniego, dr. filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki *Czechowicz.* *Kurup.* i *Maloch.* Obraz *Czechowicz.*
Lachowicki h. *Lodzin.* w orsząńskim pow. Z tych Piotr podpisał elekcyą Jana III. *Nie.*
Lachowicki h. *Steniska.* Jan syn Alexandra stolnika wielkiego 1701 i Katarzyny Gorwanienickiej, borsudczy krzemieniecki, miał za sobą *Jano* coto Maryanna Reeczyka, *Żdo* coto Maryanna Bakuwika; z pierwszą był syn Feliks, którego z Klary Rahobinskiej córka Maryanna poślubiła *Gen. Jelowickich.*
Lachowicz h. *Lodzin.* Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1800. Do tego d. Jan nakaz Stanisław August, współczesny autor w Petersburgu.
Lachowicz h. *Rudawa.* *Kurup.* i *Maloch.*
Lachowski h. *Nolca.* Tadeusz i Mikolaj udowodnił szlachectwo w sądzie z przemyskiem 1782.
Lachowski h. *Pilawa.* w sandomierniem i innych województwach. N. miał za sobą *Zofia* księżna Weroniecka około r. 1520. Wojciech wojak z wiskiej, z sejm 1581 komisarz do granic pruskich od Mazowsza. N. pojął Łukowską, córkę Jana pisarza w krakowskiego, z *Zofia* Palczowskiej urodzoną. Jan w sandomierniem 1632. Jednej komornik czeski 1674. *Nie.* — Niewiadomo czy do tych należał Sebastian Juscia, sławny kasnodzieja króla Stanisława Augusta. *Kroz.*
Lacki h. od Autorowa nasi piza, że dawniejszy przodek tego domu Zawichojki nazwany w przedk. w Moskiewie był osadzi, gdzie nazwany był Lackim, a *Dobosier* w Kazaniu Wyprowadzenie *Gońcy* przysłał, że ten Zawichojki Zacharyasz postąpił z *Polski* do Moskwy, tam był osadzi; syn jego według niektórych, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który niemogąc anielsi tyranii Oweyny opiekuna młodego cara Iwana, do Litwy się przesłał za *Zygnuntia* i króla polskiego, od którego wdał dobra Zoludek i Wysoki Dwór, był oraz leźniczym podkajm, mał wojenny i w sprawach rygarskich biegły, a Bogdana Polubinska spłodził córkę Teodorę za Eliaszem Pielgrzymowskim pisarzem lit. i synów dwóch: Grzegorza i Teodora, obadwa byli działami kawalerskimi wstawiemi, osobliwie jednak Teodor, którego krawce i przeważnie *Gońcy* z pochwałą wspomina *Konasty 1611 fol. 45 i 1613;* umarł z wielkim mająm ludzi rygarskich w Wilnie, zastawiony z Izabelli Bonarulli de Rovere, znaczej we Włoszech paroneli, córkę Katarzynę za Teodorem Pielgrzymowiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę zmadek, raduńsk, i zaleski, również wojennego mecia, ten fundował OO. Dominikanów w Wysekim Dworze; żona jego Joanna Talwosówna, jedynaczka córka Adama kaszelska żmudzkiego, z której opłódz córki Izabelli *Jano* coto za Teodłem Tępną wojewodą brzeskojaskim, *Żdo* coto za Hieronimem Chłkiewiczem starostą, mozyrykim, zostali synów trzech: Teodor, Samuel i Piotr;

list szlacheckich są umieszczone w tomie I.

FAMILIE SZLACHECKIE*)

L.

Labady. *Idaly.* w konst. r. 1750 fol. 227. *Malach.*
Labach Szeyma h. Lubicz z odmianą. Patrz Szeyma.
Lachnicki, w wilkomirskim itd. *Malach.* — Ignacy sędzią polskimi, wydał *Biografię włościanina* i przełożył tragedję Woltera „Alzacy” (Warsz. 1780).
 Ignacy Emanuel (ur. 1783 † 1830) synowiec poprzedniego, dr. filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki Czechowicz. *Kurup.* i *Malach.* Obacz Czechowicz.
Lachowicki h. Łodzień, w orszakiem pow. Z tych Piotr podpisał elekcyę Jana III. *Niec.*
Lachowicki h. Sienica. Jan syn Aleksandra stolnika wielunińskiego 1701 i Katarzyny Gorwaniec-ki, horodniczy krzemieniecki, miał za sobą *Imo* woto Maryannę Rzeczycka, *Zdo* woto Maryannę Bąkowską; z pierwszej był syn Feliks, którego z Klary Reboznickiej córka Maryanna podbiła Adolfa Jelowickiego. *Gen. Jelowickich.*
Lachowicz h. Łódzie. Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1800, to samo Wojciech (ale herb jego nie wyrażony) w sądzie z Iwowskim r. 1783. — Do tego domu zdaje się należeć Stanisław August, współczesny sutor w Petersburgu.
Lachowicz h. Radwan. *Kurup.* i *Malach.*
Lachowski h. Wąlecz. Tadeusz i Mikołaj udowodnił szlachectwo w sądzie z przemyskim 1782.
Lachowski h. Piława, w sandomierskiem i innych województwach. N. miał za sobą Zofią księżną Włoszacką około r. 1520. Wojciech wojaki z wiskit, z sejmu 1591 komisarz do granic pruskich od Marowa. N. pisał Łukowską, córka Jana pisarza z krakowskiego, z Zofią Palczowską wdowianką. Jan w sandomierskiem 1632. Jędrzej komonik czeski 1674. *Niec.* — Niewiadomo czy do tych należał Sebastian Januska, sławny kaznodzieja króla Stanisława Augusta. *Kroz.*
Lacki h. wd. Autorowa nasi pisał, że dawniejczy przodek tego domu Zawichojski narwany, z Polski wyszedłszy, w Moskwie był osiadł, gdzie narwany był Lackim, a Dubowicz w kazaniu *Wyprawienie Oboj* przydaje, że ten Zawichojski Zacharyasz posługując z Polski do Moskwy, tam był osiadł; syn jego według niektórych, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który utemogąc znieść tyranii Oweczyny opiekuna młodego cara Iwana, do Litwy się przeniósł za Zygmunta I króla polskiego, od którego wziął dobra Żoludek i Wysoki Dwór, był oraz leśniczym polskim, miał wojenny i w sprawach rycerskich biegły, z Bogdaną Polubińską spłodził córke Teodorę za Eliaszem Pelgrzymowiczem pisarzem lit. i synów dwóch: Grzegorzem i Teodorem, obadwa byli dzielnymi kawalerzami walewni, osobliwie Jolank Teodor, którego krewne i przawne dzieła z pochwałą wspominał *Konary 1811 fol. 45 i 1013*; umarli z wielkim żalem ludzi rycerskich w Wilnie, zostawiliż z Izabell Bonarilli de Rovere, szlachnej we Włoszech paronteli, córke Katarzynę za Teodorem Tyaskiewiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę żmudzką, raduńską i zosińską, również wojennego męża, ten fundował OO. Dominikanów w Wysockim Dworze; żona jego Joanna Talwosówna, jedynaczka córka Adama kasztelana żmudzkiego, z której oprócz córki Izabelli *Imo* woto za Teofilem Teyzną wojewodą brzeszczańskim, *Zdo* woto za Hieronimem Chodkiewiczem starostą mozyrskim, zostały synów trzech: Teodor, Samuel i Piotr;

*) Herby rodzin szlacheckich są umieszczone w tomie I.

FAMILIE SZLACHECKIE.*)

L.

Labady. *Idaly.* w konst. r. 1750 fol. 227. *Malach.*
Labach Szeyma h. Lubicz z odmianą. Patrz Szeyma.
Lachnicki, w wilkomirskim itd. *Malach.* — Ignacy sędzią gubernialnym, wydał *Biografię włościanina* i przełożył tragedję Woltera „Alzacy” (Warsz. 1780).
 Ignacy Emanuel (ur. 1783 † 1830) synowiec poprzedniego, dr. filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki Czechowicz. *Kurup.* i *Malach.* Obacz Czechowicz.
Lachowicki h. Łodzień, w orszakiem pow. Z tych Piotr podpisał elekcyę Jana III. *Niec.*
Lachowicki h. Sienica. Jan syn Aleksandra stolnika wielunińskiego 1701 i Katarzyny Gorwaniec-ki, horodniczy krzemieniecki, miał za sobą *Imo* woto Maryannę Rzeczycka, *Zdo* woto Maryannę Bąkowską; z pierwszej był syn Feliks, którego z Klary Reboznickiej córka Maryanna podbiła Adolfa Jelowickiego. *Gen. Jelowickich.*
Lachowicz h. Łódzie. Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1800, to samo Wojciech (ale herb jego nie wyrażony) w sądzie z Iwowskim r. 1783. — Do tego domu zdaje się należeć Stanisław August, współczesny sutor w Petersburgu.
Lachowicz h. Radwan. *Kurup.* i *Malach.*
Lachowski h. Wąlecz. Tadeusz i Mikołaj udowodnił szlachectwo w sądzie z przemyskim 1782.
Lachowski h. Piława, w sandomierskiem i innych województwach. N. miał za sobą Zofią księżną Włoszacką około r. 1520. Wojciech wojaki z wiskit, z sejmu 1591 komisarz do granic pruskich od Marowa. N. pisał Łukowską, córka Jana pisarza z krakowskiego, z Zofią Palczowską wdowianką. Jan w sandomierskiem 1632. Jędrzej komonik czeski 1674. *Niec.* — Niewiadomo czy do tych należał Sebastian Januska, sławny kaznodzieja króla Stanisława Augusta. *Kroz.*
Lacki h. wd. Autorowa nasi pisał, że dawniejczy przodek tego domu Zawichojski narwany, z Polski wyszedłszy, w Moskwie był osiadł, gdzie narwany był Lackim, a Dubowicz w kazaniu *Wyprawienie Oboj* przydaje, że ten Zawichojski Zacharyasz posługując z Polski do Moskwy, tam był osiadł; syn jego według niektórych, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który utemogąc znieść tyranii Oweczyny opiekuna młodego cara Iwana, do Litwy się przeniósł za Zygmunta I króla polskiego, od którego wziął dobra Żoludek i Wysoki Dwór, był oraz leśniczym polskim, miał wojenny i w sprawach rycerskich biegły, z Bogdaną Polubińską spłodził córke Teodorę za Eliaszem Pelgrzymowiczem pisarzem lit. i synów dwóch: Grzegorzem i Teodorem, obadwa byli dzielnymi kawalerzami walewni, osobliwie Jolank Teodor, którego krewne i przawne dzieła z pochwałą wspominał *Konary 1811 fol. 45 i 1013*; umarli z wielkim żalem ludzi rycerskich w Wilnie, zostawiliż z Izabell Bonarilli de Rovere, szlachnej we Włoszech paronteli, córke Katarzynę za Teodorem Tyaskiewiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę żmudzką, raduńską i zosińską, również wojennego męża, ten fundował OO. Dominikanów w Wysockim Dworze; żona jego Joanna Talwosówna, jedynaczka córka Adama kasztelana żmudzkiego, z której oprócz córki Izabelli *Imo* woto za Teofilem Teyzną wojewodą brzeszczańskim, *Zdo* woto za Hieronimem Chodkiewiczem starostą mozyrskim, zostały synów trzech: Teodor, Samuel i Piotr;

FAMILIE SZLA

L.

Labady. *Idaly.* w konst. r. 1750 fol. 227. *Malach.*
Labach Szeyma h. Lubicz z odmianą. Patrz Szeyma.
Lachnicki, w wilkomirskim itd. *Malach.* — Ignacy sędzią polskimi, wydał *Biografię włościanina* i przełożył tra-
 Ignacy Emanuel (ur. 1783 † 1830) synowiec poprzedniego, dr. filozofii, był redaktorem kilku czasopism w Wilnie.
Lachowicki Czechowicz. *Kurup.* i *Malach.* Obacz Czechowicz.
Lachowicki h. Łodzień, w orszakiem pow. Z tych Piotr z Lachowicki h. Sienica. Jan syn Aleksandra stolnika wielunińskiego 1701, horodniczy krzemieniecki, miał za sobą *Imo* woto M. Bąkowską; z pierwszej był syn Feliks, którego z Klary Adolfa Jelowickiego. *Gen. Jelowickich.*
Lachowicz h. Łódzie. Leon udowodnił szlachectwo w wydziale Stanów galic. r. 1800, to samo Wojciech (ale herb jego nie wyrażony) w sądzie z Iwowskim r. 1783. — Do tego domu zdaje się należeć Stanisław August, współczesny sutor w Petersburgu.
Lachowicz h. Radwan. *Kurup.* i *Malach.*
Lachowski h. Wąlecz. Tadeusz i Mikołaj udowodnił szlachectwo w sądzie z przemyskim 1782.
Lachowski h. Piława, w sandomierskiem i innych województwach. N. miał za sobą Zofią księżną Włoszacką około r. 1520. Wojciech wojaki z wiskit, z sejmu 1591 komisarz do granic pruskich od Marowa. N. pisał Łukowską, córka Jana pisarza z krakowskiego, z Zofią Palczowską wdowianką. Jan w sandomierskiem 1632. Jędrzej komonik czeski 1674. *Niec.* — Niewiadomo czy do tych należał Sebastian Januska, sławny kaznodzieja króla Stanisława Augusta. *Kroz.*
Lacki h. wd. Autorowa nasi pisał, że dawniejczy przodek tego domu Zawichojski narwany, z Polski wyszedłszy, w Moskwie był osiadł, gdzie narwany był Lackim, a Dubowicz w kazaniu *Wyprawienie Oboj* przydaje, że ten Zawichojski Zacharyasz posługując z Polski do Moskwy, tam był osiadł; syn jego według niektórych, Jan Wasil Lacki za cara Wasila Iwanowicza, tego zaś syn Iwan, który utemogąc znieść tyranii Oweczyny opiekuna młodego cara Iwana, do Litwy się przeniósł za Zygmunta I króla polskiego, od którego wziął dobra Żoludek i Wysoki Dwór, był oraz leśniczym polskim, miał wojenny i w sprawach rycerskich biegły, z Bogdaną Polubińską spłodził córke Teodorę za Eliaszem Pelgrzymowiczem pisarzem lit. i synów dwóch: Grzegorzem i Teodorem, obadwa byli dzielnymi kawalerzami walewni, osobliwie Jolank Teodor, którego krewne i przawne dzieła z pochwałą wspominał *Konary 1811 fol. 45 i 1013*; umarli z wielkim żalem ludzi rycerskich w Wilnie, zostawiliż z Izabell Bonarilli de Rovere, szlachnej we Włoszech paronteli, córke Katarzynę za Teodorem Tyaskiewiczem wojewodą nowogrodzkim i syna Jana Alfonsa starostę żmudzką, raduńską i zosińską, również wojennego męża, ten fundował OO. Dominikanów w Wysockim Dworze; żona jego Joanna Talwosówna, jedynaczka córka Adama kasztelana żmudzkiego, z której oprócz córki Izabelli *Imo* woto za Teofilem Teyzną wojewodą brzeszczańskim, *Zdo* woto za Hieronimem Chodkiewiczem starostą mozyrskim, zostały synów trzech: Teodor, Samuel i Piotr;

*) Herby rodzin szlacheckich są umieszczone w tomie I.

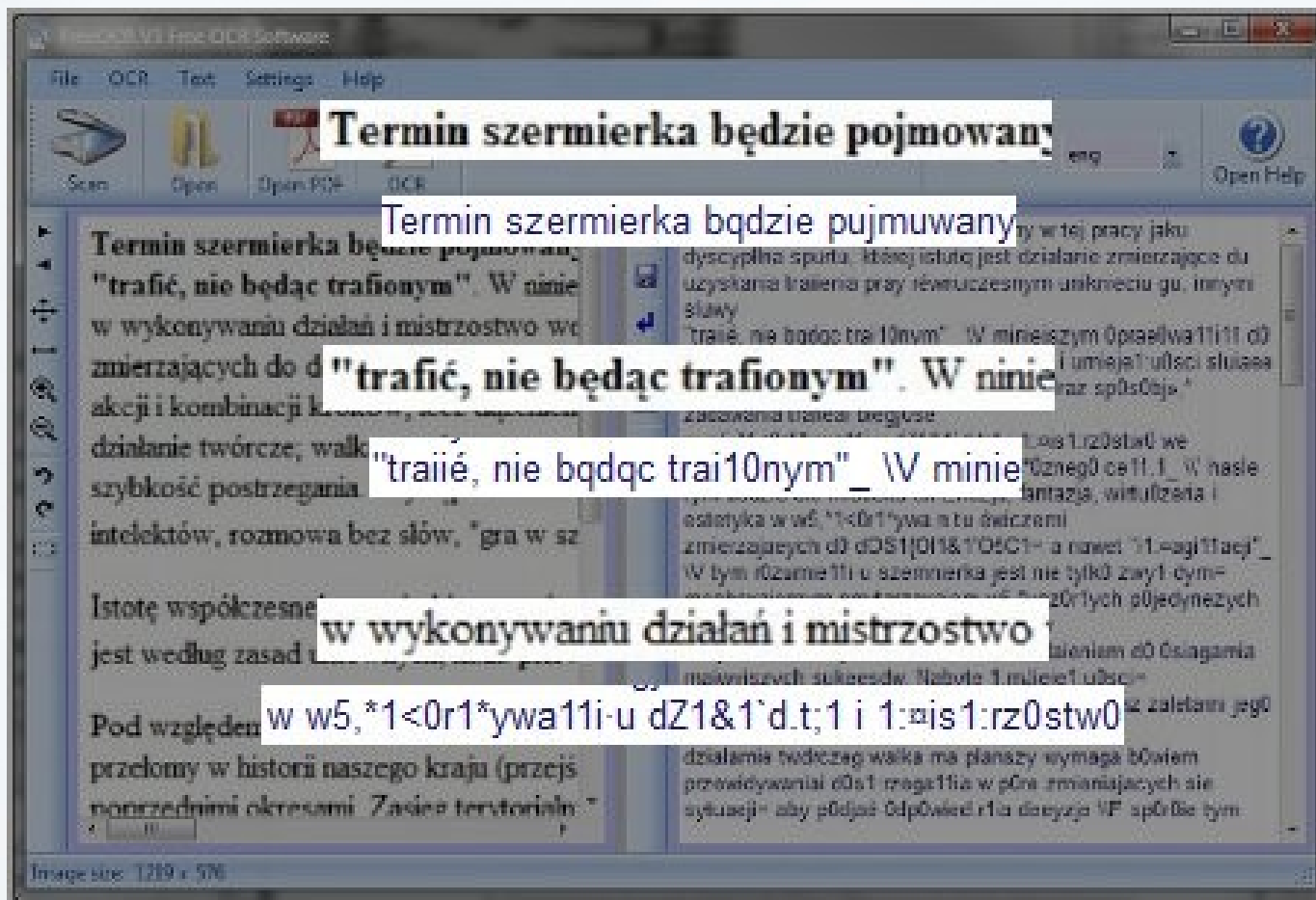
Deskew mniejszość w tomie I

Auto Manual

-1.56°

How does OCR works? (2)

- Character recognition by comparison with pattern database
 - If pattern database does not contain given character, OCR will try to match it to „something” which is similar
- Preparation of final text
 - Combining characters into words and lines according to original order of text
 - Comparing words with a dictionary, spelling correction possible



OCR – quality of results

- Quality of print and general state of the document
 - OCRs perform much better on modern printed texts
- Scanning resolution and color depth
 - >300 PPI, color depth 1 BPP
- Language of text, alphabet and font

OCR – quality of results (2)

- Text layout
 - Different in books and in newspapers
 - If layout analysis is done in a wrong way, text flow is broken
- Text formatting
 - Appropriate spacing between letters, lines and page boundaries

File formats

- Digital text can be stored as simple plain text files
 - Choose encoding wisely
 - No information about original features of the text (layout, formatting, illustrations)
- Other solutions: HTML (EPUB), PDF or DjVu
- PDF and DjVu are relatively complex file formats, allowing to:
 - Embed original image as a background for text
 - Add annotations, bookmarks
 - Optimize for web delivery

PDF

- Portable Document Format (PDF) is an open standard
- Created by Adobe Systems in 1993
- Aimed to decouple document representation from application, hardware and operating systems
- Fully featured and widely known

DjVu

- Designed to store scanned documents
 - Especially those containing a combination of text, line drawing and photographs
- Similar features as PDF, but:
 - DjVu document can be splitted into multiple files
 - Ususally smaller files
- Limited adoption in recent year
 - Text is not indexed by Google
 - Requires web browser plugin

Agenda (2)

- **A few words about digitisation hardware and software**
 - Including a quick tour through the e-learning course
- Case study: „OCR in Poland”
- Discussion

Agenda (2)

- A few words about digitisation hardware and software
 - Including a quick tour through the e-learning course
- **Case study: „OCR in Poland”**
- Discussion

OCR in Polish digital libraries

- Survey was held in September-October of 2010
- We received responses from 26 major institutions
- Survey covered 70% of resources gathered by Polish Digital Libraries Federation
- We were asking about all sort of things related to creation of full text search
 - Type and number of documents
 - Digitisation practices
 - Usage of OCR software

OCR in Polish digital libraries (2)

- 40% of objects were a subject of OCR
- No one does the correction of OCR results
- Scanning resolution between 300-600 PPI, colour depth depends on type of resource and purpose of digitisation
- FineReader and Document Express are most widely used software packages
- Only 3 institutions use training capabilities of OCR software

OCR in Polish digital libraries (3)

- Used OCR software
 - Document Express
 - ABBYY products
 - Omnipage (1 institution)
- Input formats
 - TIFF (25 institutions), DjVu (17), JPEG (5)
- Output formats
 - DjVu, PDF
 - ABBY Finereader XML (5 institutions)

Agenda (2)

- A few words about digitisation hardware and software
 - Including a quick tour through the e-learning course
- Case study: „OCR in Poland”
- **Discussion**

Discussion

- Any thoughts, questions?

Discussion

- What are scenarios for digitisation in small memory institution?
- What software is used in your library?

References

- Lessons from DRMSI course:
 - „What should I know about digitisation equipment?”
 - „What should I know about software used in digitisation process?”
 - „What is the rationale for digitisation activities in the cultural heritage domain?”