

UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA  
FACULTATEA DE ISTORIE ȘI FILOSOFIE

Teză de doctorat  
Rezumat

CORDONATOR  
CS II dr. LUMINIȚA DUMĂNESCU

DOCTORAND  
ANGELA CRISTINA LUMEZEANU

CLUJ-NAPOCA  
2019

UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA  
FACULTATEA DE ISTORIE ȘI FILOSOFIE

**INFRASTRUCTURI DIGITALE PENTRU ISTORIA SOCIALĂ.  
CONSTRUIREA BAZELOR DE DATE ISTORICE**

COORDONATOR  
CS II dr. LUMINIȚA DUMĂNESCU

DOCTORAND  
ANGELA CRISTINA LUMEZEANU

CLUJ-NAPOCA  
2019

# CUPRINS

Lista de imagini .....	4
Lista tabelelor .....	7
Introducere .....	8
Bazele de date și cercetarea istorică: Prezentare generală .....	8
1. Structura de lucru HPDT, HDG și baze de date secundare .....	11
2. Surse și metodologie. Limbaje și medii de lucru: prezentare din punct de vedere tehnic .....	13
3. Mulțumiri .....	16
Capitolul 1 .....	18
Bazele de date și cercetarea istorică: Un Instrument transformator .....	18
1. Bazele de date și schimbările din cercetarea istorică .....	20
1.1. Demographic Databases from Sweden – POPLINK and POPUM .....	23
1.2. Norwegian Historical Population Register (NHD) .....	25
1.3. The Historical Sample of Netherlands (HSN) .....	27
1.4. Integrated Microdata Series (IPUMS and NAPP) .....	28
1.5. MOSAIC Database .....	30
1.6. The Karelian database (KATIHA) .....	32
1.7. Scanian Economic Demographic Database .....	33
1.8. Antwerp - *COR database .....	34
1.9. Hungarian Historical Demographic Database .....	35
1.10. Historical Population Database of Transylvania – HPDT .....	35
2. The European Historical Population Sample Network (EHPS-Net) .....	39
2.1. Intermediate Data Structure (IDS) .....	41
3. Concluzii .....	44
Capitolul 2 .....	46
Ce este o bază de date? Informații de bază despre funcționalitatea bazelor de date.....	46
1. Baze de date concepte de bază .....	46
2. Concluzii.....	62
Capitolul 3 .....	64
Historical Population Database of Transylvania .....	64
1. Surse .....	64
2. Registrele parohiale ca surse pentru HPDT .....	65
2.1. Registrele parohiale în Transilvania .....	68

2.1.1. Registrele de botez .....	75
2.1.2. Registrele de căsătorie .....	83
2.1.3. Registrele de înmormântare .....	96
3. Construirea setului de date .....	107
3.1. Baza de date a surselor .....	107
3.1.1. Tabelul de Nașteri (Botez) .....	108
3.1.2. Tabelul de căsătorii .....	120
3.1.3. Tabelul de înmormântări .....	129
3.1.4. Tabele secundare și constrângeri .....	132
3.2. Baza de date standard .....	144
3.2.1. Procesul de inter-relaționare computerizată .....	151
3.3. Baza de date Open access .....	159
4. Concluzii .....	163
Capitolul 4 .....	165
Historical Data Grinder .....	165
1. Sursele istorice și modelul HDG .....	166
1.1. Surse primare .....	167
1.1.1. Surse de arhivă .....	167
1.1.2. Colecții de documente .....	172
1.1.3. Colecții de corespondență .....	174
1.1.4. Memorii .....	177
1.1.5. Presa .....	178
1.2. Surse secundare .....	181
1.2.1. Lucrări de referință (dicționare, enciclopedii, bibliografii, șematisme etc.) .....	181
1.2.2. Hărți, liste de localități și alte surse legate de geografia administrativă și populație .....	184
1.2.3. Literatură de specialitate (cărți, studii, articole) .....	184
1.2.4. Surse online .....	186
3. Modelul Entitate-Atribut-Valoare și HDG .....	187
4. Concluzii .....	202
Capitolul 5 .....	204
Invalizi, orfani și văduve de război. Baza de date IOVR .....	204
<b>1. Sursele ce prezintă victimele Primului Război Mondial .....</b>	<b>204</b>
<b>2. De la informație dispersată la construirea bazei de date IOVR .....</b>	<b>209</b>
<b>3. Concluzii .....</b>	<b>224</b>
Capitolul 6 .....	225
Probate Database of Transylvania .....	225

1. Introducere: istoriografie, motive și scop .....	226
2. Surse și structura PDT ( <b>versiunea 1.0.</b> ) .....	231
3. Dezvoltare pe viitor: Conscripțiile de secol XVIII și inventare de avere .....	245
4. <b>Concluzii</b> .....	250
<b>Concluzii</b> .....	252
<b>Referințe</b> .....	272
Anexa 1. Diagramă a relațiilor dintre entități HPDT .....	284
Anexa 2. Diagramă a relațiilor dintre entități HDG .....	285
Anexa 3. Diagramă a relațiilor dintre entități IOVR .....	286
Anexa 4. Diagramă a relațiilor dintre entități PDT .....	287

**Cuvinte cheie:** Transilvania, baze de date relaționale, baze de date de populație istorică, model orientat spre sursă, model orientat spre metodă, Historical Population Database of Transylvania (HPDT), Historical Data Grinder (HDG), War Invalids, Orphans and Widows Database (IOVR), Probate Database of Transylvania (PDT)

## INTRODUCERE

Prezenta lucrare își propune să ofere istoricilor din România o discuție aplicată, în profunzime, a principiilor care guvernează construirea bazelor de date, punând accentul pe bazele de date destinate cercetărilor demografice și de istorie socială. Această imagine de ansamblu va menține o abordare accesibilă pentru publicul nespecialist, fără a insista excesiv asupra aspectelor tehnice ale construirii unei baze de date; vor fi evidențiate, în schimb, aspectele structurale și conceptuale despre care istoricii nu au cunoștințe aprofundate (la fel ca majoritatea cercetătorilor din științele umaniste, în general). Abordarea acestor chestiuni este extrem de necesară în contextul actual, având în vedere faptul că în fiecare an apar tot mai multe baze de date, portaluri și colecții online de documente, sau proiecte cu o componentă digitală. În tot acest timp, cunoașterea generală referitoare la transpunerea surselor istorice în baze de date rămâne, în cel mai bun caz, la un nivel incipient, fapt care împiedică dezvoltarea acestora la potențial maxim. Tipurile de baze de date și principiile care le guvernează rămân doar parțial cunoscute, motiv pentru care sunt valorificate la doar o mică parte din potențialul lor (Erickson 2013, 134). Pentru a remedia această situație, diferite lucrări, precum cea de față, au fost dedicate unor discuții accesibile, dar totuși aprofundate, despre ceea ce presupune construirea unei baze de date, și despre modul în care folosirea unei asemenea baze de date poate transforma, în mod pozitiv, cercetările referitoare la Transilvania în domenii precum demografia, istoria socială, economică sau chiar politică.

Această lucrare a fost inspirată de două necesități: în primul rând, așa cum am arătat mai sus, s-a făcut resimțită nevoia de face cunoscută, într-o formă accesibilă și unei noi categorii de beneficiari, și anume istoricilor, metodologia construirii unei baze de date; în al doilea rând, este necesară documentarea construirii și dezvoltării bazei de date Historical Population Database of Transylvania (HPDT), cel mai complex proiect implementat până acum în România de creare a unei infrastructuri digitale pentru cercetarea istorică, socială și demografică. În ceea ce privește lucrarea de față, în prima parte voi discuta structura acesteia, apoi voi oferi o scurtă trecere în revistă, cu caracter predominant tehnic, a limbajului și modalităților de construire, comune pentru toate bazele de date documentate în această teză. Am inclus această vedere de ansamblu în introducere tocmai pentru a sublinia funcționalitatea și caracteristicile comune tuturor bazelor de date. Am ales, în acest fel, să prezint

separat aspectele pur tehnice referitoare la conceperea unei baze de date, destinate unei audiențe mai specializate, față de secțiunile cu caracter descriptiv analitic și istoric. Lucrarea este împărțită în șase capitole principale dedicate unor aspecte teoretice și conceptuale mai largi, precum și aplicării principiilor referitoare la proiectarea bazelor de date și la construirea unor diverse infrastructuri digitale.

HPDT a fost construită pe baza registrelor parohiale cu informații despre botez, logodnă, căsătorie, înmormântare, schimbarea confesiunii, confirmare, vaccinare etc. Historical Data Grinder (HDG) se bazează, de asemenea, pe surse de arhivă dar într-un format mai degrabă narativ, precum presă, corespondență, memorii, anuare ale unor instituții etc. Baza de date IOVR (Invalidi, Orfani, Văduve de Război) centralizează informații din dosarele indivizilor înregistrați pentru a primi pensie, dosare care includ o largă varietate de documente: acte de stare civilă, rapoarte despre caracteristicile personale ale soldaților înrolați, informații despre participarea la război, mărturii referitoare la dispariție sau moarte, înregistrări medicale, acte privind situația familiilor beneficiarilor și suma de bani primită, adeverințe referitoare la modificarea situației acestor beneficiari. Probate Database of Transylvania (PDT) conține registre cu testamente și documente referitoare la succesiune. Pentru mai bună înțelegere a complexității bazelor de date, fiecare capitol al lucrării începe cu descrierea amănunțită a surselor folosite la construirea acestora.

Informația referitoare la sursele istorice este completată cu literatura de specialitate privind construirea și gestionarea bazelor de date. Bazele de date HPDT, HDG, IOVR și PDT au fost construite în MySQL.

MySQL este un sistem relațional de management al bazelor de date (RDBMS) deținut de Oracle și care are la bază Structured Query Language (SQL). Este cea mai populară bază de date open acces. MySQL funcționează pe multiple sisteme de operare, inclusiv Windows, Linux și platforma Unix. În majoritatea situațiilor, MySQL este asociată cu diferite aplicații web, fiind concepută să gestioneze rapid baze de date de mari dimensiuni.

Aplicațiile web folosite pentru a crea interfața grafică sunt scrise în limbajul de programare Ruby și pe structura Ruby on Rails. Rails este o structură model-view-controller (MVC) care permite dezvoltatorului să scrie mai puține coduri obținând aceleași rezultate (spre exemplu, HPDT are 71 000 de linii de cod). Este tot un program open source.

Toate aplicațiile sunt găzduite pe un server Ubuntu folosind Ruby Version Manager (RVM). RVM permite gestionarea diferitelor versiuni instalate de Ruby și Rails pentru că fiecare aplicație are propriul pachet de funcții și programe care trebuie instalate. Legătura dintre aplicații și serverul web Apache este făcută prin intermediul Phusion Passenger, care este un server puternic pentru aplicații web.

## **CAPITOLUL 1. Bazele de date în cercetarea istorică: un instrument transformator**

Primul capitol al acestei lucrări a fost dedicat trecerii în revistă a bazelor de date existente, cu accent pe cele care sistematizează informațiile extrase din sursele istorice. Bazele de date pot veni în întâmpinarea a numeroase cerințe în cercetarea istorică, cele mai des întâlnite fiind gestionarea datelor, interconectarea înregistrărilor (record linkage) și analiza datelor.

Kees Mandemakers și Lisa Dillon (Mandemakers & Dillon 2010) au elaborat un ghid de bune practici pentru crearea bazelor de date orientate către surse (*source-orientated database*), în care au evidențiat importanța selecției tipurilor potrivite de surse din care să se extragă un set de date care să fie inclus în baza de date, recomandând de asemenea ca sursele să fie transcrise ca atare, formând de fapt o clonă digitală a documentului scris. Aceștia au scos în evidență și importanța modului de introducere a datelor, integrarea, stocarea, și diseminarea acestora. Ei consideră necesară existența unei baze de date pentru surse, a uneia pentru datele standardizate și a uneia pentru uzul public. Un alt model de bază de date aflat în uz este așa-numita bază de date orientată spre metodă (*method oriented database*). Aceasta este mai ușor de conceput și de construit, iar colectarea datelor necesită mai puțin timp. Pentru acest model de bază de date însă este absolut necesară cunoașterea de la început a aspectelor care se intenționează a fi cercetate, ba chiar și categoriile de interogări formulate.

Pentru istoricii care studiază populația, cea mai potrivită este baza de date orientată către surse. Includerea integrală a populației dintr-o regiune anume în baza de date poate lua foarte mult timp. Acesta este motivul pentru care unele baze de date, începute în urmă cu câteva decenii, sunt încă în curs de dezvoltare. Majoritatea acestor baze de date dedicate studiilor de populație și demografiei sunt centralizate online în cadrul unei rețele, European Historical Population Samples Network (EHPS-Net), create în 2011 pentru a aduce împreună cercetătorii din lumea întreagă cu scopul de a crea un format comun pentru diferite baze de date de populație (<https://ehps-net.eu/databases>). Printre acestea se poate număra și HPDT.

## **CAPITOLUL 2. Ce este o bază de date? Principii fundamentale ale funcționalității unei baze de date**

Al doilea capitol este dedicat modelelor folosite la construirea unei baze de date și modalităților de stocare și accesare a datelor. Ambele aspecte sunt relevante pentru modelele pe care le-am folosit la crearea bazelor de date pe care le-am dezvoltat. Cu ocazia cercetărilor derulate pentru elaborarea acestei teze, am constatat că există numeroase interpretări eronate și utilizări greșite ale termenului „bază de date” în special în domeniul istoric. Din acest motiv, unul dintre scopurile principale ale lucrării de față a fost să ofere istoricilor o sinteză a cunoștințelor de bază cu privire la ceea ce înseamnă o bază de date (spre deosebire de alte medii de stocare a informației în format digital), și despre modul în care sursele istorice pot fi translatate în formatul unei baze de date relaționale. Există numeroase tipuri de fișiere și obiecte care pot stoca date. Totuși, pentru ca acestea să poată fi considerate baze de date, este nevoie ca aceste fișiere să poată stoca fizic undeva nu doar informațiile, ci și legăturile dintre acestea. De asemenea, o bază de date trebuie să ofere modalități eficiente de accesare și modificare a informațiilor (adăugare, actualizare, ștergere și vizualizare). Atunci când se iau în considerare elementele unei baze de date și relația dintre acestea, apar evidente anumite modele, precum cel relațional sau cel grafic.

Modelul relațional a fost teoretizat pentru prima dată de E.F. Codd într-un articol publicat în 1970, prin care au fost stabilite o serie de reguli de bază. Conform acestora, o bază de date este ”o colecție bine organizată de informații care sunt legate într-un mod semnificativ și care pot fi accesate prin mai multe decizii logice”. Informațiile brute sunt stocate într-o manieră organizată, iar cantitatea acestora poate fi uriașă. Există un program cunoscut ca sistem de management al datelor (data management system - DBMS) care face posibilă administrarea bazei de date și servește ca interfață între cererile utilizatorilor și locul unde sunt stocate datele. Actualmente sunt disponibile numeroase asemenea DBMS, precum Microsoft Access, FileMakerPro, Oracle, Mysql, PostgreSql, DB2 etc. Aceste sisteme trebuie să întrunească următoarele caracteristici: să ofere o modalitate de a introduce, modifica și șterge informații, de a extrage informații (prin intermediul unor interogări, de exemplu), să permită restricționarea accesului la date prin crearea unor utilizatori desemnați și parole.

Baza de date grafică este un DBMS care funcționează pe baza unui model grafic de date. Acesta constă din două elemente: entități (persoană, loc, categorie, profesie etc.) numite noduri și relații. Fiecare relație reprezintă modul în care aceste entități - noduri sunt asociate unele cu altele. Acest tip de bază de date oferă o mare flexibilitate în special pentru integrarea într-un sistem coerent a unor

documente, texte, literatură etc. Baza de date grafică nu este potrivită pentru gestionarea unor cantități mari de date și interogări care acoperă întreaga bază de date.

O bază de date spațială este o colecție de informații optimizată pentru a stoca elemente definite într-un spațiu geometric. De obicei acestea sunt asociate cu elemente geografice, topografice etc. Informațiile din bazele de date spațiale sunt stocate în linii, poligoane, puncte, coordonate.

O bază de date relațională este un sistem în care obiectele sunt organizate într-o colecție de relații afișată în formă tabelară. Aceasta, numită și *schema*, are două componente principale: entitatea și relația. Entitatea este un obiect despre care se dorește stocarea de informații în baza de date. Entitățile corespund tabelelor din modelul relațional, și pot fi definite ca obiecte ușor recognoscibile: studenți, angajați, clienți, etc. Relația definește modul în care entitățile stocate în baza de date interacționează unele cu altele. În general, baza de date este organizată în mai multe tabele cu coloane și rânduri identificate printr-o cheie unică. Fiecare coloană reprezintă un atribut al unei entități. Conexiunea logică între aceste tabele este reprezentată de relație. Fiecare coloană are o valoare unică; fiecare rând are un identificator unic, numit cheie primară (*primary key*); nu pot fi două coloane identice (trebuie să aibă nume diferit); toate intrările referitoare la o coloană trebuie să fie de același tip.

Există trei tipuri de relații care definesc legătura dintre entități: 1: 1 (one to one), 1: m (one to many) și m: n (many to many). Un factor important care afectează performanța bazei de date este operațiunea de join. Funcțiile de join sunt elemente de legătură între tabele în modelul relațional, și toate informațiile pot fi extrase atâta timp cât aceste legături sunt aplicate corect. Pentru a evita redundanța în construirea unei baze de date relaționale, este folosit procesul de normalizare.

Pe parcursul dezvoltării unei baze de date este foarte importantă parcurgerea unor activități fundamentale. Jan. L Harrington le-a denumit "ciclul de viață al unui design structurat" (*structured design life cycle*) (Harrington 2009), acesta incluz stabilirea și evaluarea nevoilor, a modelului bazei de date, dezvoltarea și implementarea bazei de date. După implementare urmează o fază de testare, apoi acest ciclu este reluat până când baza de date ajunge la configurația optimă.

### Capitolul 3. Historical Population Database of Transylvania

Al treilea capitol este dedicat bazei de date ”Historical Population Database of Transylvania” (HPDT). HPDT a fost rezultatul unui proiect al cărui scop principal a fost construirea unei baze de date de populație pentru perioada 1850-1914 în Transilvania. Este capitolul cel mai consistent pentru că este prima și cea mai complexă bază de date de populație istorică pentru România. Include trei baze de date separate cu mai multe versiuni de dezvoltare fiind un model pentru alte baze de date implementate în domeniul istoriei din România.

Capitolul se deschide cu sursele folosite la crearea bazei de date. Principalele surse sunt registrele de stare civilă (botezuri, căsătorii, înmormântări) completate și păstrate de către biserică. Caracteristic pentru Transilvania este faptul ca registrele parohiale au fost scrise în mai multe limbi (română, maghiară, latină, germană etc.) cu alfabet și scrieri diferite (latin, chirilic, Kurrentschrift etc.). De asemenea registrele provin de la diverse confesiuni: ortodoxă, greco-catolică, catolică, reformată, luterană, izraelită etc. Aceasta diversitate s-a dovedit a fi o provocare pentru arhitectul bazei de date și pentru operatori deoarece baza de date a fost permanent ajustată și modificată ca să corespundă fidel surselor.

Informația din HPDT respectă cu strictețe sursele originale, urmărind modelul orientat către surse, chiar dacă operatorul a observat în surse erori de scriere, înregistrare sau chiar simple omisiuni care ar fi putut fi completate ușor prin deducție logică. Astfel, informația din baza de date respectă în întregime caracterul sursei, cu toate erorile sale potențiale sau modificări. Baza de date este realizată în engleză (inclusiv observațiile operatorului), dar informația ca atare a fost transcrisă în limba originală a registrului fără a se depărta de sursă.

Mai departe, în capitolul trei se analizează construirea setului de date și urmărește descrierea celor trei baze de date ce compun HPDT: baza de date a surselor, baza de date cu informația standardizată și baza de date open access. De la început baza de date a fost gândită să aibă o componentă de cercetare și una publică open access. Baza de date de cercetare, disponibilă numai după autentificarea utilizatorului, are o bază de date a surselor, o bază de date cu informația standardizată și un tabel cu persoane interconectate. Baza de date publică (<http://hpd.t.ro:4080>) este inclusă într-o pagină web care oferă publicului larg o perspectivă asupra dinamicii populației din Transilvania între 1850-1914. Informația prezentată în baza de date publică conține o parte consistentă din baza de date a surselor.

Toate componentele bazei de date ”Historical Population Database of Transylvania” (HPDT) sunt dezvoltate în modelul relațional al bazelor de date folosind sistemul de management al datelor MySQL,

o arhitectură open source și au o interfață grafică dezvoltată în programul Ruby on Rails de asemenea open source. Utilizatorul nu are acces direct la baza de date din MySQL dar se conectează la ea prin interfața grafică ce îi permite adăugarea, editarea și ștergerea înregistrărilor. Arhitectura urmează modelul de bune practici în construirea bazelor de date istorice. HPDT cuprinde persoane atestate în registrele parohiale de la diferite confesiuni.

Fiecare tabel detaliază un eveniment major din ceremonia religioasă: botez (*Tbirths*), logodnă (*Tbetrothals*), căsătorie (*Tmarriage*) și înmormântare (*Tdeaths*). În completare la aceste trei tabele principale sunt mai multe tabele complementare. Fiecare tabel este conectat la celălalt printr-o coloană de legătură care permite administrarea unor interogări complexe, căutarea și regăsirea datelor într-un mod logic și eliminarea redundanței. Conexiunea logică între tabele se face prin coloana de legătură numită foreign key. La aceste foreign keys se adaugă constrângeri, funcții care asigură că baza de date nu va avea înregistrări orfane (fără legături) în tabelele dependente.

Toate tabelele secundare sunt legate de unul sau toate cele 4 tabele principale printr-una sau mai multe coloane. Sunt 24 de tabele cu elemente predefinite care ajută la menținerea informației într-o formă normalizată. Tabelul *Tbirths* are în prezent 21 conexiuni (relații), *Tmarriages* are 44 de conexiuni (relații), *Tdeaths* are 20 de conexiuni (relații) și *Tengagements* are 7 conexiuni (relații). O parte din aceste conexiuni sunt referințe către același tabel. De exemplu în tabelul *Tmarriages* există 7 conexiuni (relații) către tabelul *Tdenominations* corespunzând confesiunii fiecăruia dintre participanții la eveniment: mirele, mireasa, părinții mirilor și martorii.

Pentru a putea menține un echilibru între principiile modelului de bune practici în crearea bazelor de date istorice și principiile unei baze de date relaționale au fost create tabele pentru Preoți și Moașe. Pentru a elimina redundanța, tabelul *Tpriests* înregistrează numele preotului doar o singură dată deși numele lui apare la mai multe evenimente și este legat de înregistrarea evenimentelor prin *priest\_id* care este foreign key. În același mod este construit și tabelul pentru Moașe – *Tmidwives*. Acest tabel este legat de tabelul *Tbirths* (botezuri) tot prin foreign key - *midwife\_id*.

Există un tabel care documentează sursa originală (*Tsources*) – unde apar informații legate de limba folosită, alfabetul, parohia și confesiunea, data primei și ultimei înregistrări în sursă.

În continuare al treilea capitol descrie baza de date cu informația standardizată. Pentru a avea o calitate mai bună a rezultatelor în procesul de inter relaționare computerizată informația a fost prima data standardizată. Tabele din baza de date a surselor au fost copiate într-o nouă bază numită Baza de date standard.

Pe lângă tabelele din baza surselor au fost adăugate alte câteva tabele. Cel mai important este cel a numelor (*Tnames*). Standardizarea s-a făcut pentru prenume, nume și poreclă. Un al doilea tabel este

cel pentru locație geografică, care conține toate numele localităților menționate în baza de date a surselor cu un corespondent standardizat și codat.

Un alt tabel standard este cel a ocupațiilor – *Toccupations*. Toate ocupațiile menționate în baza de date au fost adăugate în acest tabel și au fost codate folosind ”Historical International classification of occupation” HISCO.

După standardizare urmează procesul de inter-relaționare computerizată. Inter-relaționarea computerizată este procesul căutării și găsirii seturilor de date din surse multiple și care se referă la aceeași entitate și apoi unificarea acestor informații disparate. Astfel fiecare fragment de informație despre acea entitate este unificat obținându-se la final o înregistrare completă. Există diverse metode automate de a standardiza informația care se regăsește în surse care nu au un numitor comun (id, număr de identificare al persoanei (CNP), număr social de identificare etc.)

În România procesul de inter-relaționare computerizată nu fost utilizat niciodată în cercetarea istorică deoarece până în prezent nu a existat nicio bază de date de populație istorică suficient de mare încât să permită o astfel de cercetare. Construirea bazei de date HPDT a creat structura necesară pentru a încerca metode automate de inter relaționare. Programul realizat s-a bazat pe formula matematică a calculării distanței Jaro-Winkler. Această formulă măsoară asemănările dintre două șiruri de caractere (nume, localități, ocupații etc.). Pentru a putea aplica metoda automată de inter relaționare computerizată a fost creat un nou tabel – *Individuals*. Acest tabel a fost completat cu un eșantion de date extrase din cele trei tabele principale.

Baza de date open access este prezentată în ultima parte a capitolului trei. The Historical Population Database of Transylvania (HPDT) are o componentă accesibilă publicului larg (<http://hpdt.ro:4080>). Conține toate evenimentele înregistrate în baza de date a surselor privind botezuri, căsătorii și înmormântări. Nu toate coloanele din baza de date a surselor au fost afișate ci doar câmpurile relevante.

Proiectarea și construirea unei baze de date, în special utilizând modelul orientat spre sursă este o sarcină care consumă mult timp. Atunci când se lucrează cu surse complexe, informație eterogenă această sarcină devine dificilă deși rezultatele sunt foarte promițătoare pentru cercetătorii care studiază demografia istorică. Toate aceste rezultate și în special procesul de inter relaționare computerizată depind foarte mult de calitatea surselor. Transilvania are o poziție unică pentru că aici diferitele confesiuni oferă o varietate de surse cu o multitudine de câmpuri de informație. Problema apare atunci când toată această varietate trebuie integrată în modelul orientat spre sursă. În viitor ar fi interesant de văzut cum arhitectura HPDT va putea fi adaptată pentru a deveni un model pentru realitățile istorice din spațiul est european.

## Capitolul 4. Historical Data Grinder

Următorul capitol al tezei este dedicat dezvoltării unui alt model de bază de date istorică, și anume Historical Data Grinder. Meritul principal al acestei baze de date este faptul că recunoaște diferitele sisteme de cercetare ale istoricilor oferind astfel o alternativă la modelul relațional clasic.

Sursele istorice includ un număr mare de instrumente biografice și prosopografice cum sunt dicționarele, diverse liste de nume, schematisme. Din păcate, ele sunt disponibile numai în format tipărit sau în versiuni electronice care nu sunt ușor de accesat și nu încurajează abordări comparative.

Luând în considerare acești factori și nevoia unui cadru comun unde informația provenită din surse multiple poate fi stocată și legată cu alte baze de date existente, considerăm că cel mai potrivit este modelul Entitate-Atribut-Valoare (EAV). De asemenea, modelul EAV este conform cu cunoștințele legate de calculator ale istoricilor români.

Principalul avantaj al modelului EAV este flexibilitatea. Cea mai simplă formă a acestui model conține doar trei tabele pentru entitate, atribut și valoare și unul pentru tabelul de metadate. Adăugarea unei entități noi nu necesită crearea de câmpuri noi și reproiectarea bazei de date de către programator. Este o sarcină simplă ce constă în adăugarea unui rând nou în tabelul de entități, activitate ce poate fi făcută de către cercetător. În același mod sunt adăugate atributele noi în tabelul de Atribute. Formatul datelor este foarte curat, relațiile sunt foarte clare și modelul este mult mai potrivit atunci când atribute noi trebuie înregistrate în baza de date în mod regulat.

Pentru istorici, provocarea principală o reprezintă înțelegerea modelului EAV. Am creat un model Historical Data Grinder (HDG) unde informația este descompusă conform modelului EAV. Există două modalități de a înregistra informația în baza de date. În prima informația se procesează offlin: istoricul își poate organiza datele în fișiere Excel urmărind modelul formularelor din baza de date. Apoi informația este importată în masă în baza de date prin fișiere .csv. A doua modalitate este de adăuga/edita înregistrarea una câte una prin formularele bazei de date. Această operație necesită o conexiune permanentă la internet și mai mult timp pentru completare.

În structura actuală, baza de date acceptă în teorie orice tip de informație istorică, de la reconstituirea relațiilor de familie în cercuri mici sau în grupuri locale sau geografice până la reproducerea evoluției administrative cu modificările în timp (și acestea împărțite pe categorii). HDG are o capacitate sporită de acționa atât ca un agregator cât și ca o bază de date independentă.

În interfața grafică listele de înregistrări sunt prezentate într-o formă tabelară. În același timp este disponibilă descărcarea înregistrărilor în fișiere .csv și .xls pe baza unor filtre de căutare pentru a putea extrage doar informația necesară cercetării.

Modelul Entitate-Atribut-Valoare (EAV) se potrivește mult mai bine nevoilor istoricilor care folosesc surse variate și au nevoie să le integreze într-o analiză complexă. Flexibilitatea este avantajul principal oferit de modelul EAV, în timp ce provocarea principală rămâne înțelegerea mecanismelor de funcționare a modelului EAV și cum se pot unifica datele într-un model coerent care permite analize multiple și complexe.

## **CAPITOLUL 5. Invalizi, orfani și văduve de război. Baza de date IOVR**

Capitolul cinci este dedicat bazei de date Invalizi, orfani și văduve de război (IOVR). Dosarele persoanelor înregistrate pentru a primi pensie conțin o întreagă varietate de documente, cele mai importante fiind acte de stare civilă, rapoarte individuale despre soldații înrolați, informații despre participarea la război, acte doveditoare ale decesului sau dispariției, acte medicale, documente referitoare la situația familiilor beneficiarilor și la modificările survenite de-a lungul timpului, suma primită.

Baza de date a fost structurată în trei tabele principale (Invalizi, Văduve și Orfani), două tabele secundare (Copii, Evenimente) și cinci tabele auxiliare (Comitate, Locuri, Relații, Grade militare, Unități). Trei tabele sunt folosite în comun cu HPDT (Confesiune, Stare civilă și Ocupație). Mai sunt, de asemenea, trei tabele de legătură care conectează tabelul Ocupații cu tabelele principale, cu scopul de a crea o relație de tip m: n (many to many) pentru a permite selectarea multiplă. Modelul IOVR are în prezent 16 tabele și este strâns conectată cu HPDT, dar poate fi, de asemenea, separată și construită ca o bază de date independentă dacă este nevoie.

Baza de date conține formulare cu intrări pentru invalizi, văduve și orfani. În viitor se intenționează conectarea acestor date cu informații din HPDT, ceea ce va permite identificarea și urmărirea indivizilor în Transilvania pe parcursul epocii moderne.

## CAPITOLUL 6. The Probate Database of Transylvania

Capitolul șase (și ultimul) este dedicat bazei de date Probate Database of Transylvania (PDT). Construită cu un scop diferit față de Historical Population Database of Transylvania și Historical Data Grinder, PDT a fost concepută să includă informații din câteva tipuri de surse bine definite, referitoare la caracteristici sociale, profesionale și economice ale indivizilor. Deși este asemănătoare cu HPDT ca și concepție (orientată spre surse mai degrabă decât orientată spre metodă), această bază de date adoptă o abordare mai fragmentată a conținutului informațional al surselor folosite, decât cea folosită în cazul registrelor parohiale.

Scopul Probate Database of Transylvania PDT este acela de a combina informații la nivel individual provenite din mai multe tipuri de surse social-economice și juridice.

Probate Database of Transylvania este împărțită în prezent în 12 tabele și este găzduită pe același server ca și HPDT și celelalte baze de date amintite mai sus, și are în comun cu HPDT interfața grafică pentru utilizator. Ca și HPDT, PDT a fost construită în .sql, arhitectura sa fiind astfel concepută încât să se adapteze diversității surselor. În același timp, spre deosebire de HPDT, nu reproduce sursele cu fidelitate, ci le descompune în elementele principale ale acestora.

Structura Probate Database of Transylvania urmărește două direcții principale, determinate de cele două tipuri de surse pe care le încorporează. Fiecare tip de sursă are propriul tabel: *Wills\_Sources* (Surse testamentare) și *Register\_Sources* (Surse registre). Tabelul *Individuals* (Indivizi) este legat de tabelele *Civil Status* (Stare Civilă), *Ethnicity* (Etnie), *Profession* (Profesie), *Role* (Rol); un alt tabel important este *Relation* (Relații), în care figurează, codificate, relațiile dintre indivizii din sursele introduse în baza de date.

În timp ce adoptă o parte dintre principiile și liniile directoare aparținând atât modelului orientat către surse, cât și modelului Entitate-atribut-Valoare EAV, PDT se deosebește de acestea în mai multe aspecte: sunt preluate separat doar entitățile principale precum indivizii, diviziunile sau testamentele în timp ce alte secțiuni urmăresc în mai mare măsură sursele, pentru a reține marea varietate de attribute și valori conținute de acestea.

Deoarece procesul de inter relaționare computerizată (record linkage) va avea în vedere, în primul rând, indivizii și fragmente de proprietăți imobiliare (în măsura în care este posibil), informația este introdusă gradual, accentul fiind pus pe combinarea datelor despre persoanele incluse în surse, unde anume dețineau proprietăți, și care era valoarea acestora. Deși această abordare se abate de la unele

principii trasate de Mandemaakers și Dillon (2004), sarcina s-a dovedit mult mai ușoară și a consumat mai puțin timp decât reproducerea fidelă a surselor istorice în integralitatea lor, înainte de a face orice formă de standardizare.

## CONCLUZII

Dezvoltarea rapidă a tehnologiei informației (IT) a avut un impact masiv în toate domeniile științifice. În ultimele decenii, acest impact a început să fie resimțit și în domeniile de cercetare umanistă, deodată cu cristalizarea unor discipline precum "social science history" sau emergența unor periodice precum *Historical Computing* sau *Historical Methods*. După ce istoricii ca reprezentanți a unei ramuri științifice au devenit din ce în ce mai familiarizați cu folosirea computerului, foarte multă informație care înainte a fost prelucrată manual a putut fi analizată automat, ceea ce a eliberat mult timp pentru a explora o multitudine de tehnici de cercetare și orientări teoretice. Cercetarea a devenit mult mai ușoară din multe puncte de vedere, ca de exemplu cel al spațiului de stocare, modalităților de accesare și recuperare a datelor, precum și datorită posibilității de a interoga și analiza cantități mari de date în intervale de timp foarte scurte.

Astăzi nu ne-am putea imagina viața fără utilizarea calculatorului. Informația pe care o strângem de-a lungul cercetărilor este enormă și necesită o platformă pentru sistematizarea acestor date adunate, care să permită studierea lor și diseminarea către alți cercetători. Ca urmare, toată această informație trebuie organizată într-un mod care să permită cercetătorului să administreze, să prelucreze și să distribuie rezultatele. Cea mai bună soluție este un sistem de management al datelor, pe scurt o bază de date, care s-a dezvoltat la o scară din ce în ce mai larg datorită necesității tot mai mari a cercetătorilor.

De când Codd a teoretizat primul model de bază de date relațional (Codd, 1970), utilizarea bazelor de date a crescut în mod exponențial. Acum utilizarea lor face parte din activitatea noastră zilnică, schimbându-se ca urmare și modul de abordare al cercetării. Dezvoltarea unor baze de date mari de către cercetători precum și folosirea tehnicilor moderne de statistică și concentrarea pe fundamentele sociale sunt trăsături definitive ale istoriei actuale. Dezvoltarea bazelor de date a fost esențială în transformarea metodelor de cercetare.

Toată informația prezentată mai sus în cele patru baze de date românești va fi în viitor relaționată și inter-conectată. Aceste baze de date prezintă un model conceptual care tratează surse diferite, reflectând direcția și conținutul cercetării istorice actuale, în cadrul căreia informația trebuie recuperată din mai multe surse. Ceea ce construirea acestor infrastructuri digitale a arătat este faptul că heterogeneitatea și calitatea acestor surse este trăsătura principală care dictează arhitectura bazei de date, pe lângă întrebările de cercetare ale istoricilor și diferențele dintre raportul dintre costuri și beneficii în cazul orientărilor metodice (method-oriented) sau al orientării către surse (source-oriented) în construcția unei baze de date. Experiența de construcție a HPDT, HDG, IOVR și PDT a demonstrat că soluția

ideală pentru Transilvania este o remodelare a modelului clasic orientat către sursă astfel încât să se adapteze surselor prezente în acest spațiu printr-o combinație între date standardizate și date în forma lor originală.

Mai mult, experiența creării acestor infrastructuri digitale a arătat de asemenea importanța comunicării informate între istorici și arhitectul bazei de date, acest prim pas cheie în implementarea celei mai bune soluții tehnice pentru cercetare. Deși nu e de așteptat ca niciuna dintre părți să devină un specialist în domeniul celeilalte, istoricii care își doresc să includă o componentă digitală în proiectele lor de cercetare – mai ales sub forma unei baze de date – trebuie să aibă măcar o înțelegere de bază a principalelor concepte de construcție și funcționare a unei bazei de date, pentru a-și structura sursele în cea mai potrivită manieră. Deși expertiza arhitectului bazei de date în lucrul cu sursele istorice este esențială, este și de datoria istoricilor să comunice într-un limbaj cât mai clar scopurile analitice ale proiectelor lor și specificitățile surselor cu care lucrează. De asemenea, este necesar ca aceștia să aibă o privire de ansamblu asupra surselor lor, și mai ales asupra gradului acestora de heterogeneitate, înainte de a începe construcția propriu zisă a unei baze de date. Lucrarea de față a încercat astfel să se constituie într-o primă piatră de temelie pentru istoricii români interesați de construirea și lucrul cu bazele de date, concentrându-se pe cunoștințele fundamentale despre construcția infrastructurilor digitale și plecând de la exemple oferite de proiectele de acest tip recent implementate în România.

## Bibliografie selectivă

### Surse electronice:

<http://port.sas.ac.uk/mod/book/view.php?id=75&chapterid=132>).  
<http://port.sas.ac.uk/mod/book/view.php?id=75&chapterid=133>  
<http://s3.thinkaurelius.com/docs/titan/0.5.4/getting-started.html>  
<http://soc.kuleuven.be/web/staticpage/6/30/eng/819>  
<http://www.cedar.umu.se/english/ddb/public-access-databases/>  
<http://www.censusmosaic.org/data>  
<http://www.karjalatk.fi/index.php?lang=en>  
<http://www.rhd.uit.no/sources.html>  
<https://katiha.mamk.fi/>  
<https://socialhistory.org/en/hsn/index>  
<https://www.ehps-net.eu/databases>  
<https://www.ehps-net.eu/databases/antwerp-cor-database>  
<https://www.ehps-net.eu/databases/historical-population-database-transylvania-1850-1914>  
<https://www.ehps-net.eu/databases/historical-sample-netherlands>  
<https://www.ehps-net.eu/databases/hungarian-historical-demographic-database>  
<https://www.ehps-net.eu/databases/karelian-database>  
<https://www.ehps-net.eu/databases/mosaic-project>  
<https://www.ehps-net.eu/databases/norwegian-historical-population-register-1800-1964>  
<https://www.ehps-net.eu/databases/poplink>  
<https://www.ehps-net.eu/databases/popum>  
<https://www.linkedin.com/pulse/comprehensive-analysis-nosql-vs-rdbms-rassul-fazelat>  
<https://www.mysql.com>  
<https://www.ed.lu.se/databases/sedd>  
<https://www.ipums.org/>  
<https://usa.ipums.org/usa/>  
<https://international.ipums.org/international/>  
<https://www.nappdata.org/napp/>  
<https://www.ehps-net.eu/databases/scanian-economic-demographic-database>  
<https://www.ed.lu.se/databases/sedd>  
<https://www.ed.lu.se/databases/sedd/sedd-public-access>  
<https://www.studytonight.com/dbms/first-normal-form.php>  
<https://www.tutorialspoint.com/difference-between-schema-and-database-in-mysql>  
<http://dev.mysql.com/doc/refman/5.7/en/create-procedure.html>)  
<http://dev.mysql.com/doc/refman/5.7/en/select.html>  
<http://ec.europa.eu/eurostat/web/nuts/local-administrative-units>  
<http://www.mysqltutorial.org/mysql-boolean/>  
<http://www.mysqltutorial.org/mysql-foreign-key/>  
<https://dev.mysql.com/doc/refman/8.0/en/char.html>  
[https://dev.mysql.com/doc/refman/8.0/en/date-and-time-functions.html#function\\_date-format](https://dev.mysql.com/doc/refman/8.0/en/date-and-time-functions.html#function_date-format)  
<https://dev.mysql.com/doc/refman/8.0/en/date-and-time-types.html>

<https://port.sas.ac.uk/mod/book/tool/print/index.php?id=75>  
<https://www.britannica.com/topic/dispensation>  
<https://www.tutorialspoint.com/mysql/mysql-data-types.htm>

## Bibliografie

1. “Regulament pentru funcționarea Oficiului Național Invalizi, Orfani și Văduve și pentru organizarea și dezvoltarea operelor de asistență națională” / “Regulation for the Functioning of the National Office for the Invalides, Orphans, and Widows, and for the Organization and Development of the National Assistance Works”(1922). *Monitorul Oficial*, 29 aprilie: 874-901.
2. ”Historical Sample of the Netherlands (HSN)”. (2015). *Annual Report*. International Institute for Social History, Amsterdam.
3. ”Intoxicants & Early Modernity, England, 1580-1740”, Project information available at <https://www.intoxicantsproject.org/publications/database/>.
4. **Alter, G., Mandemakers K.** (2014). ”The Intermediate Data Structure (IDS) for Longitudinal Historical Microdata, version 4”, *Historical Life Course Studies*, vol. 1, 1-26, published 26.05.2014. Available at [http://www.ehps-net.eu/article/intermediate-data-structure-ids-longitudinal-historical-microdata-version-4/hislives\\_vol\\_1\\_03\\_07\\_14.pdf](http://www.ehps-net.eu/article/intermediate-data-structure-ids-longitudinal-historical-microdata-version-4/hislives_vol_1_03_07_14.pdf).
5. **Alter, G., Mandemakers, K. and Gutmann M.P.** (2009), Defining and Distributing Longitudinal Historical Data in a General Way through an Intermediate Structure, *Historical Social Research / Historische Sozialforschung*, 3 (109), 34 : 78-114.
6. **Bouchard, G.** (1998). ”The Content of Occupational Titles: What Can Be Learned from a Population Register?”. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 31(2): 75-86.
7. **Chelaru, R.D.** (2018). ”The Virtual Genealogical Archive (ArhGenVirt). A Pilot Digitization Project of the Parish and Civil Status Registers in the Bucharest and Brasov County Archives (Romania)”, *Transylvanian Review*, 27(3): 140-154.
8. **Codd, E.F.** (1970), “A Relational Model for Large Shared Data Banks”, *Communications of the ACM*, 13(6): 377–387.
9. **Codd, E.F.** (1985a), “Is Your DBMS Really Relational?”, *ComputerWorld*, 14 October.

10. **Codd**, E.F. (1985b), "Does Your DBMS Run By the Rules?". *ComputerWorld*, 21 October.
11. **Codd**, E.F. (1990), *The Relational Model for Database Management*, 2nd ed. Addison Wesley Publishing Company.
12. **Crăciun**, B., **Holom** E.C., **Popovici** V. (2015). „Historical Population Database of Transylvania: Methodology Employed in the Selection of Settlements and Micro Zones of Interest”, *Romanian Journal of Population Studies*, 2(9): 17-30.
13. **Dillon**, L., **Amorevieta-Gentil**, M., **Caron**, M., **Lewis**, C., **Guay-Giroux**, A., **Desjardins**, B. & **Gagnon**, A. (2018). "The Programme de recherche en demographie historique: past, present and future developments in family reconstitution", *The History of the Family*, 23(1): 20-53.
14. **Dinu** V. & **Nadkarni** P. (2007). "Guidelines for the Effective Use of Entity–Attribute–Value Modeling for Biomedical Databases", *International Journal of Medical Informatics*, 76(11-12), 769-779.
15. **Edvinsson**, S. (2000). "The Demographic Data Base at Umeå University - a resource for historical studies". In P.H. Hall, R. McCaa, & G. Thorvaldsen, *Handbook of International Historical Microdata for Population Research*. Minnesota Population Center.
16. **Farhat**, Y., **Martins**, M. & **Swanson**, D.A. (2014). *Methods of Demographic Analysis*, Springer.
17. **Feneșan**, C., ed. (1986). *Izvoare de demografie istorică*, Vol. 1, *Secolul al XVIII-lea. Transilvania*. București: Direcția generală a arhivelor statului din RSR.
18. **Gavrilă**, I. (2011). *Baze de date istorice. Cercetarea istorică asistată de calculator. Rezultate statistice ale comparării automate a bazelor de date istorice*. Oscar Print, București.
19. **Glavatskaya**, E., **Thorvaldsen**, G., **Fertig**, G. & **Szolysek**, M. (2019). *Nominative Data in Demographic Research in the East and the West*, Ekaterinburg: Ural University Press.
20. **Harrington**, J.L. (2009). *Relational Database Design and Implementation: clearly explained*, 3rd Ed., Morgan Kaufmann Publishers-Elsevier, Burlington MA.
21. **Harrington**, J.L. (2010). *SQL Clearly Explained*, Third Edition, Morgan Kaufmann Publishers-Elsevier, Burlington MA.
22. **Helgertz**, J. (2014). "Introduction to SEDD, IDS and STATA", Lecture presented at Lund Summer School Course.
23. **Holom**, E.C., **Crăciun**, B. (2014). "Invalidi, orfani, văduve. Consecințe sociale ale „Marelui Război” în Transilvania. Studiu de caz"/ Invalids, Orphans and Widows: The Social Consequences of the "Great War" in Transylvania. Case Study". In I. Bolovan, G. Cojocaru, O.M. Tămaș (Eds). *Primul Război Mondial. Perspectivă Istorică și istoriografică/World War I*.

- A Historical and Historiographical Perspective*. Cluj Napoca: Academia Română/Centrul de Studii Transilvane, Presa Universitară Clujeană, 559-568.
24. **Holom, E.C., Sorescu-Iudean O., Hărăguș M.** (2018), "Beyond the Visible Pattern: Historical Particularities, Development, and Age at First Marriage", *The History of the Family*, 23(2): 329-358.
  25. **Mandemakers, K. & Dillon, L.** (2004). "Best practices with large databases on historical populations". *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 37 (1): 34-38.
  26. **Mandemakers, K., et al.** (2018). "HSN standardized, HISCO-coded and classified occupational titles, release 2 2018.01" (IISG Amsterdam 2018). Available at <https://datasets.iisg.amsterdam/dataset.xhtml?persistentId=hdl:10622/MUZMAL>.
  27. **Matthijs, K., & Moreels, S.** (2010). "The Antwerp COR\*-database: a unique Flemish source for historical-demographic research". *The History of the Family*. 15(1): 109-115.
  28. **Maudsley, E. & Munch T.** (1993). *Computing for historians. An introductory guide*. Manchester University Press, Manchester and New York.
  29. **Moldovan, L.** (1958). "Registre confesionale de stare civilă din Transilvania". *Revista Arhivelor*. 1: 159-185.
  30. **Moldovan, L.** (1968). "Registrele parohiale din Transilvania", *Revista de statistică*, XVII, 11: 52-60.
  31. **Nadkarni, P., & Marengo, L.** (2007), "Database Architectures for Neuroscience Applications", *Methods in Molecular Biology*, vol. 401: *Neuroinformatics*, 37–52.
  32. **Nadkarni, P., Marengo, L., Chen, R., Skoufos, E., Shepherd, G., Miller, P.** (1999), "Organization of heterogeneous scientific data using the EAV/CR representation", *Journal of the American Medical Informatics Association*, 6 (Nov / Dec): 478-493.
  33. **Popovici, V.**, (2015), "Parish registers from Transylvania – sources for the History of Medicine (late 18th – early 20th centuries)", *Acta Medico-Historica Adriatica*, 2(13): 287-302.
  34. **Räisä, J., & Lojonen, M.** (2014). "The modernization, migration and archiving of a research register", in *DLM 2014 [Recurso eletrônico]: making the information governance landscape in Europe / 7th Triennial Conference of the DLM Forum*; ed. José Borbinha, Zoltán Szatucsek, Seamus Ross. - Livro eletrônico. - Lisboa : Biblioteca Nacional de Portugal, Available at [http://purl.pt/26107/1/DLM2014\\_PDF/13%20%20The%20modernization,%20migration%20and%20archiving%20of%20a%20research%20register.pdf](http://purl.pt/26107/1/DLM2014_PDF/13%20%20The%20modernization,%20migration%20and%20archiving%20of%20a%20research%20register.pdf)

35. **Rassat, S., Ruggiu, F. J.** (2019). "Big data, 3D models and digital heritage as instruments for demographic and historical studies, the example of the 'Charleville' project", *In Situ – Revue de Patrimoines*, Issue 39, Available at <https://journals.openedition.org/insitu/22296#text>.
36. **Shorter, E.** (1971). *The historians and the computer*. Englewood Cliffs.
37. **Sigmirean, C.** (2000), *Istoria formării intelectualității românești din Transilvania și Banat în epoca modernă*, Cluj-Napoca: Presa Universitară Clujeană.
38. **Sigmirean, C.** (2007), *Intelectualitatea ecleziastică. Preoții Blajului (1806-1948)*, Târgu-Mureș: Editura Universității 'Petru Maior'.
39. **Sigmirean, C.** (2011), *Formarea elitelor militare ale Imperiului Austro-Ungar. Studenți transilvăneni la Academia Militară „Ludovika” din Budapesta*, Sibiu: Astra Museum.
40. **Sorescu-Iudean, O.** (2017). „The Permeability of Borders: Drafting Romanian Last Wills in the Seat of Sibiu during the Eighteenth Century”. *Banatica*, 27: 463-484.
41. **Sorescu-Iudean, O.** (2019). “Mai mult decât suma părților. O viziune integrativă asupra inegalităților sociale, de venit și avere în Sibiul secolului al XVIII-lea prin Probate Database of Transylvania”. Conference Presentation at the National Conference of the Centre for Population Studies, „Studii de Populație. Provocări și perspective”, 1-2 July 2019, Cluj-Napoca.
42. **Sorescu-Iudean, O., Lumezeanu, A.** (2018). “Un model digital pentru surprinderea succesiunii în Transilvania premodernă”, Conference Presentation at the National Congress of Romanian Historians, 29 August – 1 September 2018, Iași.
43. **Sumathi, S. & Esakkirajan, S.** (2010). *Fundamentals of Relational Database Management Systems*, Springer, Berlin.
44. **Szolysek, M., & Gruber, S.** (2016). “Mosaic: recovering surviving census records and reconstructing the familial history of Europe”. *The History of the Family*. 21(1): 38-60.
45. **Teorey, T.J., Lightstone, S.S., Nadeau, T., Jagadish, H.V.** (2011), *Database Modeling and Design: Logical Design*, Fifth Edition, Morgan Kaufmann Publishers-Elsevier, Burlington MA.
46. **Thorvaldsen, G.** (2000). “The Norwegian Historical Data Centre”. In P. H. Hall, R McCaa, & G. Thorvaldsen, *Handbook of International Historical Microdata for Population Research*. Minnesota Population Center.
47. **Thorvaldsen, G.** (2011). “Using NAPP Census Data to Construct the Historical Population Register for Norway”. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 11(1): 37-47.

48. **Van Leeuwen, M.H.D, Maas, I., Miles, A.** (2002). *HISCO. Historical International Standard Classification of Occupations*. Leuven: Leuven University Press.
49. **Varga, R.** (2017). "Romans 1by1 v.1.1. New developments in the study of Roman population". *Digital Classics Online* 3(2): 44-59.
50. **Varga, R., Pázsint, A.I., Boda, I., Deac D.A.** (2018). "Romans 1by1. Overview of a research project". *Digital Classics Online*, 4(2): 37-63.
51. **Willigan, J. D., Lynch, K. A.** (1982). *Sources and Methods of Historical Demography: Studies in Social Discontinuity*, Academic Press.
52. **Winchester, I.** (1992). "What every historian needs to know about record linkage for the microcomputer era". *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 25(4):149-165.
53. **Wisselgren, M., Edvinsson, S., Berggren, M., & Larsson, M.** (2014). „Testing Methods of Record Linkage on Swedish Censuses". *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 47(3): 138-151.