

# Rezumatul fazei de execuție

---

Așa cum s-a precizat în capitolul 2, obiectivul principal al acestei faze de execuție constă în studiul din punct de vedere interdisciplinar al stadiului actual de dezvoltare al domeniului, conceptelor și structurilor. Acest studiu se va realiza pe două direcții, prima legată de domeniul arhivistic, iar a doua legată de domeniul IT.

Din punct de vedere arhivistic s-au urmărit următoarele:

1. Identificarea scopului domeniului arhivelor.
2. Prezentarea unui scurt istoric al arhivelor.
3. Evidențierea noii paradigme spre care se îndreaptă arhivele, precum și stadiul actual al arhivelor în România.
4. Studiul sistemului arhivistic românesc. S-a urmărit studierea istoriei arhivelor în țara noastră, a instituțiilor de arhivă existente precum și a cadrului legislativ actual. Prin acest studiu s-a urmărit totodată evidențierea modului de ordonare și inventariere a arhivelor, precum și descrierea sistemului de evidență a fondurilor și colecțiilor.
5. Studiul arhivelor în societatea informației precum și a problemei informatizării arhivelor românești scoțându-se în evidență perspectiva informatizării arhivelor clujene.
6. Digitalizarea în contextul arhivelor. S-a făcut un studiu de caz cu privire la tipurile de cataloage digitale în spațiul germanic.

Din punct de vedere al domeniului IT s-au studiat următoarele aspecte:

1. *Ontologii*

Pentru a facilita regăsirea informațiilor relevante semantic pentru un anumit domeniu de interes este necesară elaborarea unor ontologii de domeniu care să permită îmbogățirea semantică a datelor structurate provenite din diferite surse arhivistice. Pentru aceasta este necesară înțelegerea conceptului de ontologie, a rolului ei, precum și identificarea limbajelor de reprezentare și interogare corespunzătoare. În ceea ce privește limbajele de reprezentare a ontologiilor s-au studiat: RDF, OWL și SWRL.

În continuare au fost studiate principalele limbaje de interogare a ontologiilor și anume SPARQL, SeRQL, RDQL și XsRQL. În final, accentul s-a pus pe câteva unelte existente pentru reprezentarea și reasoningul ontologiilor, unelte care vor putea fi utilizate în fazele următoare ale proiectului: Protege, Jena, Pellet și Sesame.

## 2. *Adnotarea semantică a textului*

Regăsirea informațiilor relevante semantic nu este posibilă fără adnotarea semantică a documentelor. Pentru aceasta s-au studiat aspectele generale ale adnotării semantice a textului, precum și două abordări existente. Prima abordare facilitează adnotarea semantică automată prin indexarea documentelor, iar a doua abordare permite regăsirea informațiilor pe baza ontologiilor prin adaptarea unui model vector-spațiu. În final, s-a urmărit identificarea unor unelte existente care permit adnotarea semantică a textului, accentul punându-se pe Gate, o arhitectură generală pentru *text engineering*.

## 3. *Managementul datelor semantice*

Având în vedere creșterea utilizării ontologiilor în diverse aplicații, nevoia de a furniza suport pentru acestea în sistemele de management al bazelor de date a devenit mai concretă și urgentă. De aceea este necesar să se identifice atât modalități de stocare a datelor ontologice în baze de date cât și mijloace de interogare semantică. Pentru aceasta au fost studiate câteva abordări prin care se urmărește managementul datelor ontologice în cadrul bazelor de date relaționale obișnuite și în cadrul unor baze de date hibride (relațional-XML). Totodată a fost abordată și tema învățării ontologiilor din bazele de date relaționale, pentru aceasta studiindu-se câteva abordări. În final, s-au studiat două unelte existente, Jena și KAON2, care pot fi utilizate în managementul datelor semantice.