

## Raport sinteză

### Programul Operațional Capacitate Administrativă 2014-2020

CONRENA - „Consolidarea Sistemului Statistic Național și modernizarea proceselor de producție statistică pentru efectuarea recensămintelor naționale”

Beneficiar: Lider de parteneriat: Institutul Național de Statistică

Cod SIPOCA 598/ Cod MySMIS 127577/ Nr. contract finanțare 430/17.09.2019

## Raport sinteză privind

### Activitatea A7. Introducerea de instrumente moderne pentru producerea de statistici oficiale - componenta GIS

Conform Cod SIPOCA 598/Cod MySMIS 127577/Nr. contract de finanțare 430/17.09.2019, Activitatea A7. „Introducerea de instrumente moderne pentru producerea de statistici oficiale” a început în luna septembrie 2019, având mai multe componente, printre care și cele legate de GIS pentru Recensământul general Agricol 2020 (RGA2020) și pentru Recensământul Populației și al Locuințelor 2021 (RPL2021) (compontă cu termen de finalizare februarie 2023).

România a organizat și pregătit Recensământul general Agricol 2020 (RGA2020) Recensământul Populației și al Locuințelor 2021 (RPL2021), luând în considerare prevederile Regulamentelor europene pentru fiecare dintre cele două recensăminte.

Activitatea A7 - componenta GIS a jucat un rol foarte important în pregătirea fiecăruia dintre cele două recensăminte, precum și în partea dedicată diseminării rezultatelor la nivel de grid cu suprafața de 1 km<sup>2</sup>.

Activitatea 7 - componenta GIS a avut, la rândul său, mai multe subcomponente:

- dezvoltarea aplicației WebGIS pentru actualizarea și culegerea datelor necesare;
- actualizarea datelor de intrare pentru RGA2020, dar și pentru RPL2021 (nomenclatorul stradal: denumiri de artere și numere administrative aferente arterelor, pentru localizarea locuințelor);
- sectorizarea pentru RPL2021 (Output 7C);
- pregătirea datelor și verificarea funcționării lor pe aplicația de recenzare CAPI;
- verificarea și corectarea datelor cu privire la adrese culese în cadrul recensământului în etapele CAWI/CAPI.



Toate subcomponentele de mai sus au fost realizate de specialiștii din cadrul echipei de implementare în strânsă colaborare cu experții Băncii Mondiale (BM), sub cerința RAS “**Recomandări și bune practici pentru implementarea unui sistem de management al datelor pentru datele geospațiale și sectorizarea efectivă din cadrul RPL2021**” / *“Recommendations and best practices for implementing a data management system for PHC2021 geo-spatial data and the actual sectorization.* Componenta de sectorizare face parte din livrabilele din cadrul Acordului de servicii de asistență tehnică rambursabile (RAS) privind Consolidarea Sistemului Statistic Național (P167217). Proiectul a fost implementat de Institutul Național de Statistică (INS) cu sprijinul Băncii Mondiale (BM).

Această activitate s-a desfășurat pe parcursul a mai mult de doi ani, în mai multe iterații, care au urmat procesul de pregătire și implementare al RPL2021 și amânat cu 15 luni (incluzând reprogramarea în 2022 a RPL din cauza pandemiei de COVID19, pe baza deciziei CCRPL2021 și conform **OUG 19/2020** privind organizarea și efectuarea RPL cu modificările și completările ulterioare). Motivele amânării au constat în îmbunătățirea metodologiei, având în vedere situația reală a bazelor de date (locuințe și drumuri) și îmbunătățirea acestora, având în vedere modul în care RPL2021 a aplicat sectorizarea în timpul operațiunilor de colectare a datelor în teren și ținând cont de rezultatele RPL runda 2021. Îmbunătățirile au abordat deficiențele datelor și necesitatea adaptării metodologiei de sectorizare în așa fel încât să poată prelucra datele care nu îndeplinesc cerințele specificate pentru un proces automatizat (prezentat în continuare în acest raport) observate în setul de date final înainte de a începe RPL (care a avut îmbunătățiri semnificative față de primul set de date testat, dar, totuși, a avut încă limitări, inerente unui proces dinamic de actualizare).

Pentru realizarea cu succes a RPL2021 a fost necesar, într-o prima fază, să se realizeze actualizarea nomenclatorului stradal care a stat la baza întregului recensământ al Populației și al Locuințelor 2021.

S-a creat și implementat o aplicație WebGIS ce a oferit un instrument specialiștilor INS și DTS, care a ajutat în activitățile de culegere/validare/stocare a datelor geospațiale referitoare la nomenclatorul stradal și la numerelor administrative, pentru pregătirea infrastructurii care a stat la baza recensămintelor (general agricol și al populației și locuințelor) din runda 2020. Aplicația va fi utilă și la alte activități ulterioare, în perioada intercenzitară, având în vedere evoluția localităților și a nevoii continue de actualizare a acestor seturi de date.



Principalele contribuții ale echipei INS, la realizarea lucrărilor din activitatea A7 componenta GIS pentru:

I) RGA2020, au vizat:

- *evaluare date GIS și WebGIS, evaluare coordonate GIS din baza de date APIA;*
- *analiza, verificarea, validarea coordonatelor GIS colectate la RGA2020 și imputarea datelor necollectate, din alte surse;*
- *analiza și tratarea cazurilor speciale de localizare a exploatațiilor agricole pentru care au fost colectate date RGA2020 în vederea asigurării coordonatelor GIS și a încadrării pe griduri statistice;*
- *verificare și testare aplicație determinare griduri pe baza coordonatelor GIS atribuite exploatațiilor agricole (colectate sau imputate) pe serverul Băncii Mondiale; testare și tratare cazuri speciale;*
- *verificarea corectitudinii și completitudinii codurilor grid de 1km<sup>2</sup> alocate exploatațiilor agricole prin utilizarea aplicației privind geo-referențierea exploatațiilor agricole în unități statistice de tip grid, în mai multe iterații, în funcție de fluxul de actualizare a aplicațiilor pe baza observațiilor furnizate în urma testărilor; comparații coduri INSPIRE pentru localizarea corectă a exploatațiilor agricole;*
- *constituirea bazei de date a exploatațiilor agricole cu coordonate griduri (nivel de județ). Conform Regulamentului (UE) 2018/1091, în setul de variabile obligatorii care se transmit la Eurostat, pentru fiecare exploatație agricolă trebuie inclus și codul celulei rețelei de unități statistice Inspire de 1 km<sup>2</sup> pentru utilizare paneuropeană corespunzător localizării fermei. Pentru alocarea acestui cod, fiecare exploatație agricolă a trebuit geo-referențiată, fiindu-i alocate coordonate GIS (latitudine și longitudine) aferente localizării.*

II) pentru RPL2021, au vizat:

a) *Dezvoltarea aplicației WEBGIS pentru actualizarea și culegerea datelor necesare:*

- *analiza și identificarea funcționalităților aplicației WEBGIS și definirea caietului de sarcini;*



- *evaluarea propunerilor tehnice depuse la procedura „SERVICII DE DEZVOLTARE SOFTWARE PENTRU STATISTICI (dezvoltare statistici georeferențiate - aplicație WebGIS)”*
- *analiza, verificarea și pregătirea datelor ce au fost încărcate în baza de date a aplicației WebGIS*
- *Analiza, testarea/verificarea aplicației WebGIS creată în cadrul proiectului;*
- *realizarea de instruirii cu specialiștii DTS pentru prezentarea aplicației WebGIS și a procedurii de lucru de actualizare a datelor prin intermediul aplicației;*
- *asistență tehnică acordată DTS-urilor ( mesaje scrise și telefonice) cu privire la utilizarea aplicației WebGIS cât și cu privire la implementarea procedurii de actualizarea a datelor de intrare (activitatea realizată pe toată perioada derulării activității A7);*
- *creare de scripturi SQL (bd postgresql) pentru generarea rapoartelor de activitate pentru utilizatorii WebGIS;*
- *moduri de rezolvare a situațiilor particulare identificate în procesul de actualizare a nomenclatorului stradal și a numerelor administrative - cu ajutorul aplicației WebGIS;*

*b) Actualizarea datelor de intrare (nomenclatorul stradal: denumiri de artere și numere administrative aferente arterelor):*

- *analiza și identificarea erorilor din datelor existente care conțin informații cu privire la străzi în vederea întocmirii unui nomenclator stradal pentru recensământul 2021;*
- *stabilirea etapelor de lucru pentru corectarea /actualizarea în detaliu a datelor necesare pentru realizarea recensământului 2021;*
- *creare de scripturi python pentru corectarea automată și actualizarea datelor geospațiale existente;*
- *verificarea, actualizarea prin procedee automatizate a bazei de date geospațiale a nomenclatorului stradal care a stat la baza RPL;*
- *dezvoltare metodologii de lucru pentru pentru verificarea și actualizarea nomenclatorului stradal. Procedură folosită de direcțiile teritoriale de statistică;*

- îmbunătățirea fluxului de lucru pentru actualizarea/corectarea nomenclatorului de artere atunci când a fost cazul sau când s-au întâlnit situații particulare;
- analiza geospațială a datelor existente și identificarea datelor pe baza cărora se pot stabili UAT-urile greu accesibile și clasificarea acestora;
- definirea localităților dispersate și calculul indicelui de dispersie;
- asistență tehnică oferită direcțiilor teritoriale în procesul de actualizare a nomenclatorului stradal.
- creare script pentru generare rapoarte activitate utilizatori antrenați în procesul de actualizare a nomenclatorului stradal.

Sectorizarea în contextul unui recensământ al populației și al locuințelor se referă la procesul de creare a unor sectoare ale zonei de recensământ, cu un număr similar de unități de recensământ (gospodării, locuințe) în fiecare dintre aceste sectoare.

Pentru RPL2021, aceasta înseamnă că întreaga zonă de recensământ, care este definită plecând de la *un cadru al zonei intravilane, excluzând zona extravilană*, trebuia să fie segmentată în zone geografice excluzive reciproc, cu un număr țintă aproximativ de unități de *locuințe* (și anume, de 100 sau 600), cu o abatere dorită de +/- 20%.

Metodologia și sectorizarea propriu-zisă dezvoltate în cadrul proiectului vor fi folosite în continuarea activităților INS, chiar și după finalizarea RPL2021.

Activitatea de Sectorizare pentru RPL2021 a avut la bază trei etape:

- dezvoltarea metodologiei de sectorizare;
- dezvoltarea pentru automatizarea sectorizării;
- realizarea efectivă a sectorizării.

Principalele contribuții ale echipei INS la realizarea lucrărilor din activitatea A7 componenta GIS pentru RPL, în colaborare cu experții Băncii Mondiale, au vizat:

*c) Dezvoltarea metodologiei de sectorizare:*

- analiza și pregătirea criteriilor care au stat la baza realizării metodologiei de sectorizare;
- pregătirea materialelor privind metodologia de sectorizare;
- realizarea manuală și semiautomată (folosind scripturi Python pentru procesări intermediare) a sectorizării în etapa de pilot a RPL, folosind metodologia dezvoltată;



- Generarea nomenclatorului de artere și lista numerelor administrative pentru RPL Proba;
- Analiza sectorizării pe zonele test urban și rural, realizată de către Banca Mondială;
- Generare nomenclator artere și listă numere administrative RPL Probă;
- Procedura de lucru - verificare / actualizare date RPL Probă.

*d) Dezvoltarea pentru automatizarea sectorizării:*

- participare activă alături de BM la dezvoltarea conceptului automatizării sectorizării;
- participarea activă în verificările la fiecare etapă a algoritmului folosit pentru automatizarea sectorizării;
- participarea activă în cadrul ședințelor tehnice din perioada de dezvoltare al algoritmului și a scriptului de sectorizare, alături de BM;
- verificări periodice atât a rulării scriptului de sectorizare cât și verificări ale rezultatelor pentru a se asigura îndeplinirea cerințelor minime ale metodologiei de sectorizare.

*e) Realizarea efectivă a sectorizării:*

- rularea scriptului de sectorizare la nivel de localități pentru obținerea sectorizării efective;
- verificarea rezultatelor, dacă a fost cazul corectarea posibilelor probleme în seturile de date și refacerea sectorizării în cazurile identificate cu probleme;

Echipa GIS a realizat, anterior începerii etapei CAPI, pregătirea datelor în vederea testării și testarea propriu-zisă a componentei spațiale a aplicației de recensare.

*f) Pregătirea datelor și verificarea funcționării lor pe aplicația de recensare CAPI:*

- Corecții și re/alocare adrese pentru datele CAWI;
- Transpunerea situației autorecensare din etapa CAWI în hărțile suport pentru etapa CAPI și generarea hărților;
- S-au produs datele de intrare: imagini de fundal raster TPK cât și setul de date aferent adreselor în format tabelar ce a inclus sistemul de coordonate WGS84 EPSG 4326 folosit în aplicația de recensare;
- Analiza și verificare sarcini CAPI din punct de vedere GIS.

În perioada RPL2021 cât și în perioada imediat următoare recensământului, echipa GIS alături de specialiștii IT au realizat verificarea și corectarea datelor cu privire la adrese



culese în cadrul recensământului în etapele CAWI/CAPI. Au fost identificate în cadrul acestei activități erori în seturile de date (în general erori umane la momentul completării formularelor) la nivel de adresă.

*g) Verificarea și corectarea datelor colectate CAWI/CAPI:*

- au fost analizate și corectate din punct de vedere al completitudinii și corectitudinii adreselor de localizare;
- pregătirea sectoarelor pentru cercetarea post-recensământ;
- a fost creată o procedură pentru corecții manuale de către DTS folosind aplicația opensource QGIS.

În final, în cadrul activității A7 s-a realizat prelucrarea finală a datelor RPL2021 la care componenta GIS a avut o contribuție majoră, precum și amplasarea datelor corespunzătoare unui set de indicatori principali pe Grid-uri cu suprafața de 1 km<sup>2</sup>, în vederea transmiterii lor către Eurostat și a utilizării în diseminările ulterioare - naționale și internaționale.

**Concluzie:** Prin derularea celor mai sus descrise, activitatea A7 - componenta GIS a fost implementată cu succes, conform prevederilor contractului de finanțare.

Întocmit,

Aprobat,