

Док. Class/05/NL/03 - EN



2005-->ред.1.1??--->---->2008-->ред.2!!

Довідник з імплементації

## **Зміст:**

**Вступ**

**Скорочення**

**A: Старт проекту**

*A1. Відправний пункт: інструменти від Євростат*

**B: Підготовка**

*B1. Перший крок: розробка плану імплементації*

*B2. Розробка національної версії*

*B3. Таблиці відповідності та перехідні коди*

*B4. Оцінка змін та створення дослідницького середовища*

*B5. Кінцеві терміни*

**C: Джерела**

*C1. Зовнішні джерела*

*C2. Внутрішні джерела*

**D. Засоби для спостереження**

*D1 Використання індексів*

*D2 Системи кодування за допомогою комп'ютерів*

*D3 Автоматичне кодування шляхом розпізнавання тексту як приклад (АКРТ)*

**E: Теоретичні питання (щодо здійснення спостереження)**

*E1 Визначення основної діяльності в теорії і на практиці*

*E2 Використання даних про виробництво для присвоєння нових кодів*

*E3 Використання даних структурних досліджень для присвоєння нових кодів*

*E4 Присвоєння нових кодів окремим одиницям*

*E5 Присвоєння нових кодів на рівні груп*

*E6 Присвоєння нових кодів великим одиницям*

*E7 Будування ймовірнісних кореляційних матриць*

**F: Реалізація шляхом проведення дослідження**

*F1: Підготовка до дослідження*

*F2: Проведення дослідження*

*F3: Контроль якості*

*F4: Перекодування Реєстру підприємств*

*F5: Ставлення до корегувань*

**G: Подальші заходи**

**Додаток 1: перелік членів Робочої групи з імплементації**

## Вступ

У січні 2008 року NACE (Nomenclature statistique des Activités économiques dans la Communauté Européenne - Статистична класифікація видів економічної діяльності Європейського співтовариства ) набуває нової версії Ред. 2, щоб мати можливість відображати розвиток економіки. Більшість країн ЄС мають національні класифікації, з яких можна здобути NACE. У зв'язку зі зміною NACE національні класифікації, також будуть змінюватись на більш детальному рівні, щоб відтворити національні потреби. Оскільки різні країни можуть вносити різні зміни, то і обсяг роботи також може відрізнятись.

Починаючи з лютого 2005 року Робоча група з імплементації (РГІ; перелік учасників дивіться у додатку 1) спробувала сформулювати деякі рекомендації державам Співтовариства з імплементації NACE Ред. 2 у Реєстр підприємств і статистичні дані. Цей документ є частиною зазначених рекомендацій і фокусується на питаннях, що стосуються статистичних реєстрів підприємств. Цей документ є узагальненням передового досвіду, отже, читачі можуть робити висновки з приводу того, що може підходити у їх ситуації.

Перше, що було зроблено – відправлено опитування для всіх національних інститутів статистики, щоб мати можливість обговорити питання, пов'язані з імплементацією і важливі для всіх країн.

На квітневому засіданні були обговорені результати цього опитування. І хоч між державами-членами є багато відмінностей, здається, відносно багатьох проблем, з якими має справу Робоча група, досягнуто консенсусу.

На липневому засіданні Ханс ван Хуфф [Hans van Hooff] представив першу версію цього документа. Відчувалось, що він потребує доопрацювання. Завдяки додатковій інформації від Стіва Вейла [Steve Vale] (Керівництво з імплементації і протистояння з досвідом Канади в імплементації Північноамериканської системи класифікації промисловості, 1997 р. і здійснення і ретроспективного аналізу) документ було доповнено.

У нас склалося враження, що цей перелік залишається незавершеним. Тому має тривати обговорення загальних проблем, чітке визначення взаємних проблем, що повинно привести до:

- Перерахування положень, які необхідно враховувати в процесі імплементації
- Визначення моментів, коли можна, чи необхідно надати допомогу новим державам-членам Співтовариства
- Встановлення кінцевих термінів для всіх країн

## Скорочення

АКРТ: Автоматичне кодування шляхом розпізнавання тексту

ТОДК: Телефонне опитування за допомогою комп'ютеру

КПВЕД: Класифікатор продуктів за видами економічної діяльності

ЗРП : Загальний реєстр підприємств

ІКТ: Інформаційно-комунікаційні технології

ОВД: Одиниця виду діяльності

NACE: Nomenclature statistique des Activités économiques dans la Communauté Européenne (Статистична класифікація видів економічної діяльності Європейського співтовариства)

СКП: Стандартна класифікація промисловості (тут: національна версія NACE)

РГІ: Робоча група з імплементації

## **A: Старт проекту**

### **A1. Відправний пункт: інструменти від Євростат**

Структуру NACE ред. 2 буде завершено до засідання робочої групи NACE/CRA у вересні 2005 року. Після чого її буде представлено на розгляд до SPC (Комітет з соціального захисту), після цього – Раді ЄС та Парламенту. Фактично, "остаточним" він буде після опублікування в "Official Journal", що очікується у жовтні 2006 року.

Згідно з "Календарем заходів Операція-2007" остаточний варіант пояснювальних записок буде в січні 2006 року (проект вже доступний через "Ramon"). Після чого, буде здійснено їх переклад на французьку та німецьку мови. До цих пір неясно незрозуміло, чи будуть вони перекладені на всі інші 17 офіційних мов, і якщо так, то коли. Можливо знадобиться переклад національних інститутів статистики, що, звичайно, буде додатковою роботою. Євростат з'ясовуватиме, які мови є пріоритетними. Країни, які самі складають національну версію пояснювальних записок, можуть бути менш пріоритетними, ніж ті країни, які мають чекати на переклад оригінальних пояснювальних записок рідною мовою. Пізніше перші країни можуть використати національні версії як основу для "офіційних" перекладів.

Щодо таблиць відповідності (старі-нові і нові-старі) попередні версії існували вже у лютому 2005 року, і вони будуть переглядатися ще до вересня 2005 року. Вони дуже важливі для всіх країн.

## **B: Підготовка**

### **B1. Перший крок: розробка плану імплементації**

Операція 2008 року на рівні окремих країн буде широкомасштабним проектом. Сім основних елементів такі:

- Розробка національної класифікації
- Впровадження у реєстр підприємств
- Впровадження у статистичні збірники і результати, у тому числі соціологічні спостереження
- Впровадження CRA / Prodcorn (в деяких країнах після розробки національної версії)
- Спілкування з користувачами
- Контакти з адміністративними джерелами із забезпеченням кодів видів діяльності для реєстрів підприємств
- Перегляд і затвердження нормативно-правових актів

### **B2. Розробка національної версії**

Хоч NACE – це рівень для звітування Євростату, у більшості країн можуть створюватись СКП, яка є більш конкретною, ніж NACE. Причинами можуть бути статистичний попит у країні, правові умови, або складність видів діяльності тощо. Для розробки власних національних СКП необхідні деякі передумови:

а) Національна класифікація повинна відображати NACE ред 2.

б) Повинні бути розроблені критерії для визначення, які п'ять цифр (підкласів) будуть прийняті в СКП. Кожен розподіл дорогий і вимагає поділу ресурсів. Рішення приймається лише у разі, якщо застосовуються статистично значущі, або адміністративні критерії. Важливо визначити, чи створюється класифікація лише для використання у статистичних цілях, або ж вона буде призначена і для адміністративного використання. Інші критерії можуть стосуватися розміру та

кількості одиниць у підкласі, однорідності діяльності, що класифікуються у підкласі, інтенсивності використання її за межами НІС, потреби користувачів тощо.

в) Якщо СКП буде призначений для адміністративного використання, має бути розроблений перелік причин його введення в адміністративні системи.

г) Мають бути розроблені і затверджені структура, індекси та пояснювальні записки.

д) Засоби кодування та інші системи мають бути модифіковані у відповідності до змін у СКП.

### **В3. Таблиці відповідності та перехідні коди**

Як тільки буде розроблено нову СКП, повинні бути створені таблиці відповідності між старою і новою класифікацією. Можливі такі варіанти:

- 1 до 1: старий код може бути безпосередньо перенесений до нової класифікації
- n до 1: старі коди, які мають бути переведені у один новий код, також можуть бути безпосередньо перенесені, але у такому разі потрібні будуть таблиці відповідності для того, щоб у майбутньому повертати нові одиниці у старі коди.
- 1 до n: не зрозуміло, до якого нового коду слід віднести підприємство.
- n до m: певна кількість старих кодів має бути переведена у певну кількість нових кодів.

Кожний окремий зв'язок між старим і новим кодом СКП може бути виражений перехідним кодом, який є зв'язком між старим і новим СКП, і який дозволяє здійснювати порівняння між старим і новим кодом, щоб будувати часові ряди. З такого перехідного коду можна вилучити обидві версії.

### **В4. Оцінка змін та створення дослідницького середовища**

Коли розробку NACE ред. 2 і СКП закінчено, і готові таблиці відповідності, можна побачити вплив зміни класифікації. Зміни можуть бути обговорені у НІС.

Можна здійснити оцінку щодо кількості одиниць, які можуть бути перенесені автоматично, і тих, що потребують перевірки.

Часто необхідною передумовою є створення дослідницького середовища. Оскільки уся інформація щодо перекодування повинна зберігатися, та існувати база даних по узгодженню інформації, а також засіб обробки для присвоєння кодів.

### **В5: Кінцеві терміни (Стів Вейл)**

| План імплементації                                                                                                                                        | Якомога скоріше    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|
| Структура NACE готова                                                                                                                                     | Вересень 2005 року |
| Структура національної СКП готова, одразу після неї є пояснювальні записки, індекси, засоби кодування, таблиці перетворення тощо.                         | Березень 2006 року |
| Плани імплементації у Реєстри, джерела (у НІС, зовнішні), що будуть використовуватись, необхідні зміни у спостереженнях, де є підходящими ймовірні моделі | Квітень 2006 року  |
| Зміни до бази даних реєстру                                                                                                                               | Липень 2006 року   |

|                                          |                                         |
|------------------------------------------|-----------------------------------------|
| Інформація, зібрана для нової СКП        | жовтень 2007 року                       |
| Імплементация кодів у реєстр підприємств | січень 2008 року                        |
| Подвійне кодування у реєстрі підприємств | Від січня 2008 року до грудня 2009 року |

## С: Джерела

У багатьох країнах велику роль в оновленні реєстру підприємств відіграють (адміністративні) джерела. Вони також можуть відігравати важливу роль в імплементації NACE ред. 2.

### С1. Зовнішні джерела (Стів Вейл)

Через існуючу залежність від зовнішніх реєстрів (адміністративних, таких як податкових, соціального забезпечення, торгово-промислових палат), або через відсутність можливостей, часто буває необхідним використовувати, а отже, і намагатися адаптувати ці зовнішні джерела до статистичних потреб. Сюди відносяться:

- Вчасне попередження: що дасть змогу джерелам запроваджувати зміни у зручний час
- Надання інструментарію, наприклад, пояснювальних записок, індексів, схем переходу (кореляційних матриць) зовнішнім джерелам;
- Надання допомоги в адаптації комп'ютерних систем та засобів кодування
- Розповсюдження інформації серед співробітників зовнішніх джерел, що експлуатують систему, присвоюють коди, або здійснюють перевірки (презентації, довідники, навчання).

Краще за все, коли, адміністративні джерела переходять на нову класифікацію **у той самий час**, що і реєстри підприємств НІС

Інші "зовнішні" джерела можуть бути важливим джерелом інформації:

- Торгово-промислові палати, податки (ПДВ, заробітні плати тощо), соціальне забезпечення.
- Головні організації, довідники фірм, або телефонів.
- Мережа Інтернет може бути цікавим джерелом як для окремих підприємств, так і для більш агрегованого рівня.

### С2: Внутрішні джерела

Внутрішні джерела, так само як зовнішні, можуть бути важливими для запобігання небажаних підходів до економіки. Звичайно, внутрішні джерела повинні бути узгоджені в імплементації та використанні NACE ред. 2.

- Яка інформація вже є у НІС? Він вже може володіти достатньою інформацією про підприємства в межах своїх систем, і навіть якщо вона трохи застаріла, вона може бути достатньою.
- Інформація з профільного аналізу великих і складних підприємств.
- Існуючих національних (5-ти значних) кодів СКП може бути достатньо для класифікації за NACE ред. 2.

- Може статися так, що автоматична система кодування, яка використовується, реєструє описання видів діяльності одиниці. Реалістичного описання діяльності може бути достатньо, щоб здійснити поділ за NACE ред. 2. Але для цього їй не повинно бути більше 5 років. У разі використання описання діяльності слід застосовувати оцінку достовірності. Низька достовірність означає, що підприємство підлягає наближенню (Статистичне управління Канади).

## **D: Засоби для спостереження**

Засоби мають використовуватись по мірі можливості . Деякі з них зазначені у цьому документі.

### **D1. Використання індексів**

Взагалі, використання структурно стандартизованої інформації дозволяє використовувати індекси. Перелік описання видів діяльності допомагає присвоювати кодів. Таким чином, при тлумаченні наявної інформації про діяльність, або оцінці цієї інформації із запитів, індекси можуть багато у чому допомогти всім країнам.

Держави-члени Співтовариства повинні самі розробляти індекси на рівні національних СКП. Це означає, що деякі країни повинні робити однакові завдання, особливо якщо вони спілкуються однією і тією ж мовою. Ми пропонуємо, щоб Євростат (РГІ) узгоджував їх допомогу один одному.

Для наведення довідок про індекси можна застосовувати пошукову систему.

### **D2. Системи кодування за допомогою комп'ютерів**

Як бачимо з опитування РГІ, розісланого державам-членам, не багато країн використовують засоби кодування, або інші передові системи. На даний момент здається, що АКРТ від Канади є найбільш високопродуктивною системою.

Поряд з тим, що показники з пов'язаними з ними пошуковими системами, іноді, описується як засіб кодування, існує два основних типи систем.

По-перше, є системи, засновані на *лінгвістичному проектуванні*. У такому разі текстові описання пов'язані між собою, визнаючи слова, які можуть бути у різному порядку тощо. Такий підхід залежний від мови, тому за наявності 20 мов не дуже підходить для мети нашої РГІ.

По-друге, існують системи, які використовують порівняння рядків. Вони працюють на основі збереження описання і кодів. Якщо нові описання нараховують більше рядків, ніж попередньо встановлене значення, порівняння приймається і код використовується. Є можливість встановити рівні нарахування, щоб надати різні варіанти вибору між якістю кодів і кількістю описань. Ці системи працюватимуть на всіх мовах до тих пір, поки індекс підлягає перекладу.

### **D3. Автоматичне кодування шляхом розпізнавання тексту як приклад (АКРТ)**

АКРТ – приклад використання системи порівняння рядків. В ході опитування, проведеного у лютому, було зроблено висновок, що АКРТ може зробити цінний вклад в імплементацію NACE ред. 2.

Як було сказано вище, перекодувати всі необхідні одиниці шляхом розсилки опитувань – величезне завдання. Через обмеження можливостей інститутів статистики та їх джерел і навантаження на підприємців, слід використовувати якомога більше наявної інформації.

АКРТ присвоює описанням коди. Отже, мають бути описання діяльності і вони повинні бути якомога правильними. Для того, щоб була можливість здійснити кодування, необхідно застосовувати технологію синтаксичного аналізу. Синтаксичний аналіз означає, що неопрацьований текст респондента, що вводиться, адаптується до тексту бази даних АКРТ для того, щоб його можна було порівняти і присвоїти коди.

АКРТ вимагає контексту, передусім, *бібліографічної бази даних*. Цей контекст – можливо, не лише щодо діяльності, але й, наприклад, щодо професії, категорії СРА тощо – створюється АКРТ шляхом використанням специфікацій і даних, надані користувачами у вигляді простих текстових файлів. Чим більшою буде база даних, тим більше можливостей має АКРТ. Саме така інформація необхідно АКРТ для того, щоб стандартизувати введення тексту і присвоювати коди. Також потрібні *файл для перетворення коду і перекодування і достатньо місця на диску*. Крім того, можна встановити рівні нарахування між якістю кодування і кількістю описань, що були закодовані.

### Висновки щодо АКРТ

- АКРТ потребує багато баз даних. Цю потребу можна виконати лише тоді, коли застосовується реєстр з описами (Інститутом статистики або джерелами), або усі одиниці досліджуються шляхом опитування з відкритими питаннями.
- Для співставлення необхідно, щоб База даних використовувала повні пояснювальні записки власної СКП країни
- АКРТ може використовуватись для інших класифікацій, якщо створено відповідні індекси;

Незважаючи на те, що АКРТ може бути хорошим засобом для кодування одиниць, на нашу думку, не варто зобов'язувати країни використовувати таку систему. Деякі країни не зможуть відповідати стандартам (розробка бази даних вимагає багато роботи), інші вже використовують індекси та засіб кодування, які не зовсім відповідають стандартам АКРТ. Крім того, акцент має бути на впровадженні нових кодів, а не на розробці нових інструментів.

## **Е: Теоретичні питання (під час здійснення спостереження)**

### **Е1 Визначення основної діяльності в теорії і на практиці (Арто Лутіо [Arto Luhtio])**

Як згадувалось у вступі до NACE, основний вид діяльності статистичних одиниць має визначатися відповідно до максимальної частки загальної доданої вартості цієї одиниці. Якщо це не є можливим, тоді можуть використовуватись замітники. Деякі замітники розглядаються в цьому документі. Чітких правил щодо пріоритету використання різних заміників немає. Використання певного замітника залежить від низки проблем, таких як інформація про статистичну одиницю, що розглядається, рентабельність і практика країн. Хоч основна ситуація така ж: основний вид діяльності – і другорядні види діяльності, де це доречно – усіх статистичних одиниць, що записані в реєстри підприємств, повинні бути закодовані відповідно до NACE ред.2 починаючи з 1 січня 2008 року. На практиці шляхів реалізації цього завдання – велике розмаїття.

Додана вартість найчастіше може використовуватись для груп підприємств, можливо, для підприємств, але навряд чи для місцевих одиниць. Додана вартість може бути використана для визначення основного виду діяльності, усіченої групи підприємств таким чином: річні

товарообіги резидентів з членів групи множиться на індекс доданої вартості відповідного виду діяльності. Помножені товарообіги додають на 2-значному рівні NACE і діяльність, яка робить найбільший внесок у загально додану вартість визначається як основний вид діяльності цієї групи.

Деякі види діяльності, такі як страхування та фінансове посередництво не мають товарообігу у традиційному розумінні. Основна діяльність групи, члени якої беруть активну участь у цих сферах, можуть бути визначені на основі зайнятості.

## **E2 Використання даних про виробництво для присвоєння нових кодів (Хоакім Вайсброд [Joachim Weisbrod])**

Основна проблема імплементації нової класифікації полягає в ефективному і надійному перекодуванні діяльності статистичної одиниці. Існує безліч доступних методів, суміш яких використовуватиметься в залежності від особливих обставин в різних галузях. Оскільки інформація для перекодування основного виду діяльності є обмеженою, а перекодування має здійснюватись, як статистична інформація стане доступною, повинні ефективно використовуватись інші наявні джерела інформації.

Можливим джерелом інформації є дослідження виробництва згідно з переліком PRODCOM, або його національною версією. Звичайно, це джерело інформації може бути використано тільки для одиниць дослідження PRODCOM, тобто NACE ред. 1.1. розділи C-E, втім, виробничий сектор як і раніше складає велику частину промисловості. У Німеччині збираються дані для місцевих одиниць, до речі, 80% підприємств мають лише одну місцеву одиницю і один вид діяльності одиниці.

Статистика PRODCOM готує дані про виробництво (обсяг і вартість) відповідно до загальноєвропейського узгодженого переліку продуктів. Кодування цього переліку безпосередньо пов'язане з NACE і CPA, оскільки, перші чотири цифри коду PRODCOM співпадають з відповідним класом NACE і першими шістьма цифрами у коді CPA. Одиниці зобов'язані повідомляти про своє виробництво відповідно до найбільш докладних рубрик дослідження PRODCOM. Так отримуються обсяг і вартість виробництва, які можуть бути зведені до рівня класу NACE шляхом агрегації.

Якщо новий CPA вже імплементовано у дослідження PRODCOM, вартість виробництва нових класів NACE ред. 2 можна легко об'єднати. Якщо ні, то вартість виробництва спочатку необхідно перевести у нову систему кодування. Тому дуже важливо, що перелік PRODCOM 2007 року вже подвійно закодований відповідно до старої і нової класифікації.

Основний вид діяльності визначається за допомогою методу "зверху-вниз". Щоб отримати першу вказівку на новий основний вид діяльності, цей метод можна застосовувати до вартості виробництва місцевих одиниць та підприємств.

Теоретично основна діяльність повинна визначатися відповідно до максимальної частки загальної доданої вартості, на другому місці – визначення за кількістю найманих працівників, або товарообігом. Ця інформація не доступна, коли NACE ред. 2 імплементується у реєстр. За наявності додаткової інформації та припущень, чиста вартість виробництва та / або кількість найманих працівників може бути оцінена для нових класів NACE ред. 2 до застосування методу "зверху-вниз".

Статистика Виробництва охоплює лише розділи C-E NACE ред.. 1.1., втім, це все ще велика частина статистики підприємств. Цей метод не можна використовувати механічно, але можна у поєднанні з іншими методами, такими як автоматичне перекодування та індивідуальні дослідження. Тим не менше, це хороший метод, тому що більшу частину роботи можна зробити за допомогою комп'ютера.

### **Е3 Використання даних структурних досліджень для присвоєння нових кодів (Еммануель Раулін [Emmanuel Raulin])**

Використання даних структурних досліджень для визначення головних і другорядних видів діяльності в реєстрі підприємств постійно збільшується, хоча деякі країни вважають, що це може викликати зсув результатів в реєстрі. Тим не менше, небезпеки зсуву немає у тих випадках, коли досліджуються всі одиниці вище певного порогу. Крім того, щодо використання даних вибіркового дослідження, деякі країни (наприклад, Канада) переглянули свою думку і вирішили, що вигоди переважають над негативними наслідками спричиненого зсуву, і почали їх використовувати.

Ці дані також можуть бути використані для серії щорічних статистичних даних ретроспективного аналізу. Простіше кажучи, існує два способи: "макро-метод" і "мікро-метод". Макро-метод полягає у безпосередньому перетворенні серій у новій класифікації. Мікро-метод стосується індивідуальної інформації та основної діяльності. Перевага макро-методу полягає у тому, що перерахувати ряд даних у новій класифікації буде легко. Вирішальною змінною може бути додана вартість, або кількість співробітників. Мікро-метод пропонується застосовувати для обробки майбутніх змін класифікації.

Вважається, що необхідно мати подвійне кодування принаймні для двох років поспіль, при цьому лише другий рік використовується для проведення ретроспективних досліджень.

Подвійне кодування вимагає спостереження за галузями у проміжній класифікації. У такому разі види діяльності можна наводити як у старій, так і у новій класифікації. Таким чином, потрібен лише один розподіл товарообігу. Для вторинної діяльності використовуються перехідні матриці, як окремо, так і як "середні".

### **Е4. Присвоєння нових кодів окремим одиницям (мікро-рівень)**

Класифікації мають підлягати не лише підприємства. Це також стосується місцевих одиниць і груп підприємств, а також (локальних) одиниць виду діяльності (ОВД / ЛОВД), якщо вони окремо враховуються в реєстрі.

На нашу думку, пріоритетним має бути регулярне оновлення даних на рівні окремої одиниці.

Перекодування може розпочатися з різних одиниць у різних країнах. Деякі країни (наприклад, Великобританія) розпочали перекодування на рівні локальних одиниць, а потім визначили підприємства і групи підприємств за допомогою підходу "знизу вгору". Інші країни починають на рівні підприємств і перекодовують локальні одиниці після підприємств.

На рівні одиниць може бути визначена одна основна діяльність, або одна основна чи більше другорядних видів діяльності.

Для випадків "1 до n" і "n до m" оцінку слід проводити, якщо ці одиниці є досить важливими, щоб витратити багато зусиль для отримання ідеального коду основного виду діяльності на самостійному рівні. Якщо ні, то нові коди можна відрахувати.

### **Е5. Присвоєння нових кодів на рівні груп (макро-рівень)**

Для того, щоб попередити запитання від досліджень, інформації може бути не достатньо, з іншого боку, неможливо відправити опитування занадто великій кількості одиниць. Таким чином, необхідно знайти рішення.

Можна зробити оцінку кількості одиниць, яким будуть призначені нові коди, за:

- дослідженнями. Ґрунтуючись на досвіді, або за допомогою окремо розглянутих і досліджених великих підприємств можна згадатися, як можна здійснити поділ. Інша

можливість полягає у вибіркового дослідженні в областях від "1 до n" або від "n до m". З цього виду можуть бути розроблені ймовірнісної моделі.

- ймовірнісними моделями (див. також E7). Цікаво, яких результатів можна досягнути за допомогою цих методів.

У Великобританії використовуються ймовірнісні таблиці кореляції. Є кілька варіантів, в залежності від того, з якою метою вони використовуються. Вони також можуть бути різними для підприємств та локальних одиниць.

Недоліком цих методів є те, що окремим одиницям неможливо призначити точні коди. Тому залишається коефіцієнт невизначеності, який зростатиме, якщо цей метод використовується дуже часто. Наша порада – використовувати ці моделі обережно і, якщо необхідно, тільки в статистично менш відповідних областях, або для невеликих одиниць.

В результаті на самостійному рівні, багато з цих (невеликих) одиниць можуть мати неправильні коди, але агрегати буде більш-менш правильними. Тому що вони, як правило, є локальними, і вплив на статистичні дані буде відносно низьким.

## **Е6. Присвоєння нових кодів великим одиницям**

Великі одиниці піддаються вимірюванню набагато частіше, ніж дрібні. У великих одиницях часто існує деякий особистий контакт, наприклад, на рівні управління рахунками. Тут особисті контакти і високий рівень оновлення дозволяє присвоювати коди на підставі фактичної інформації про діяльність (див. також C2).

Для більших одиниць, які не беруть участь в управлінні рахунками, перекодування можна здійснити шляхом введення додаткового або зміненого питання в опитуванні.

## **Е7. Будівництво ймовірнісних кореляційних матриць (Стів Вейл)**

Для ймовірного перекодування може бути необхідним використання кореляційної матриці. Слід втілювати такі заходи:

1. Перекодувати всі одиниці, для яких необхідна інформація, і робити це з прийнятним ступенем вірогідності
2. Внести старі і нові коди цих одиниць у таблиці і пов'язати їх
3. Видалити неприпустимі комбінації (що визначені у довідковій таблиці, що надається експертами з класифікації)
4. Розрахувати відсотки на основі даних, що залишилися (зважаючи на те, коли підрахунки особливо низькі)
5. Брати до уваги будь-які інші дані, наприклад, кореляція, яка залежить від розміру, або інша ознака одиниць
6. Створити кореляційну матрицю на основі вище зазначеного. Це може бути проста матриця, або більш складна на основі кількох змінних.

Матриця має бути перевірена, щоб можна було гарантувати обґрунтованість результатів. Слід пам'ятати, що кореляційні ймовірності можуть з плином часу змінюватися.

## **Е: Реалізація шляхом проведення дослідження**

### **Е1. Підготовка до дослідження**

Сукупність повинна будуватись з урахуванням нових членів та ігнорування тих, що зникли.

Повинні бути розроблені опитування. Загальне опитування, що застосовується до всіх видів діяльності, може попросити від респондента багато роботи у наданні необхідної інформації. З іншого боку, можна застосувати засіб кодування, щоб визначити код, заснований на цій інформації.

Опитування із закритими категоріями значно полегшує збір інформації, але не обов'язково приводить до оптимальної відповіді. Тому, у опитування необхідно перевіряти за допомогою лабораторних досліджень.

У цьому випадку замість засобу кодування ефективною може бути оптична система зчитування.

У Великобританії використовуються як відкриті, так і закриті питання: широке відкрите питання, щоб визначити широкий сектор, потім закриті питання, щоб визначити точно вид діяльності. Іноді необхідні обидва для кодування описи. Наприклад, якщо підприємство пише: "дерев'яні двері" закриті питання використовується для визначення, чи стосуються вони виробництва, чи продажів.

## **F2. Проведення дослідження**

Опитування повинні бути відправлені до усіх обраних одиниць, щодо яких неможливо отримати інформацію з інших джерел. Необхідно виробити стратегію поведінки з тими, хто не відповів на питання опитування. Сюди входить рішення щодо кількості нагадувань, засобу, що використовувався (пошта, телефон, електронна пошта) тощо. Одиниці, з якими неможливо зв'язатись, або ті, що відмовили, можуть бути класифіковані за моделей на основі ймовірності. Щоб запобігти не отриманню відповіді у листі слід наголосити на підвищеному значенні для респондента.

Звичайно, велике значення може мати Інтернет. Деякі аргументи:

- Це дешевше;
- Це знижує адміністративне тягар;
- Легше обробляти інформацію.

Звичайно, результат може відрізнятись, тому що він залежить від рівня проникнення Інтернету в економіку держав-членів.

## **F3. Контроль якості**

Інформація від досліджень має бути повною та правильною. Особливо там, де респонденти повинні виокремити інформацію, необхідно відзначити, що ця інформація є повною. З іншого боку, необхідно контролювати коди, які присвоюються вручну.

З обох причин фахівець повинен повторно розглянути вибірку дослідження сукупності з метою встановлення її якості. Це має відбуватись як регулярна перевірка.

## **F4. Перекодування Реєстру підприємств**

Після того, як було розглянуто усі одиниці, що підлягають перетворенню, стара інформація про ці одиниці у реєстрі має бути адаптована. Деякі факти до відома:

- Нова система кодування має бути своєчасно введена у РП.
- РП повинен мати обидві робочі версії NACE протягом одного або більше років. Звичайно, це для реєстру це має бути можливим. Подвійне кодування є необхідною умовою для статистичного відстеження та успішної імплементації
- Через час, необхідний для проведення досліджень (2 роки для багатьох національних інститутів статистики), деякі одиниці можуть припинити свою діяльність і що важливо,

деякі нові могли бути зареєстровані. В ході дослідження сукупності цей факт має бути врахований.

- З одиницями, які зареєстровані після завершення дослідження, слід зв'язатись розглянути їх теоретично. Без інформації код слід присвоювати на основі досвіду, або моделей. Слід враховувати межі якості.
- Повинні бути проведені перевірки якості коду. Зроблені корегування, спричинені технічною несправністю.

## **F5. Ставлення до корегувань**

Хоча число одиниць, які підлягають дослідженню, повинно бути зведене до мінімуму, це все одно буде призводити до корегування кодів поряд із зміною кодів, які описані у схеми перехідних кодів.

Через те, що зміни у NACE і СКП будуть проходити в областях, які, можливо, не можуть бути описані оптимальним шляхом, очікується, що відносно велика кількість опитувань викличе коригування. Через те, що тільки частина реєстру досліджена, створюється зсув.

Існують два способи поводження з виправленнями. Як уже згадувалось для статистики і національних рахунків важливо мати можливість ретроспективного аналізу і побудови часових рядів. Тому для ігнорування поправок можуть бути аргументи. Їм можуть дати код, знаючи, що він не правильний, і вчасно адаптувати статистику. Так, що жодних екстремальних наслідків не настане.

Більш оптимальний шлях – впровадити поправки. В цьому випадку повинні бути зроблені не лише схеми переходів, але *схеми корекції типу вхід-вихід*, щоб попередити статистику. Важливість може бути оцінена з урахуванням розміру та важливості одиниці. Звичайно, це може справити зазнає надзвичайний вплив на статистичні області.

Якщо ми допустимо поліпшення якості і не досліджуватимемо кореляції "1 до 1" і "n до 1" ми введемо зсув у реєстр (одиниць буде лише рухатися у класах "1 до n" та "n до m", а не в середину цих класів) .

Обмеження потужності, ресурси тощо можуть перешкодити поліпшенню якості.

## **G. Подальші заходи**

В якості подальших заходів необхідно здійснити такі дії:

- перевірка достовірності нових кодів у ЗРП;
- можливо: вдосконалення схем переходу та виправлення;
- дослідження, чи надають джерела ЗРП коректну інформацію;
- адаптація областей статистики і вибірки;
- національні СКП слід оцінювати, дивлячись на результати досліджень. Відсутні елементи можуть бути додані в індекси, пояснювальні записки можуть бути змінені. Також може бути прийнято рішення адаптувати 5-и значну структуру, засновану на результатах досліджень;
- засоби кодування та інші системи типізації повинні бути знову адаптовані;
- Реєстри, що використовують СКП, повинні бути в курсі всіх нових змін;
- оцінка: має бути надана інформація про результати дослідження.

### **Додаток 1: Члени РГГ**

Від Євростат:

Alice Zopré

Arto Luhtio

Leila Anupold

Michael Mietzner

Petra Sneijders

Isabelle Remond-Tiedrez

Paul Konijn

Від ЄЦБ

Heinz Christian Dieden

Від Національних інститутів статистики

Michel Blanc (Fr)

Michel Lacroix (Fr)

Thierry Lacroix (Fr)

Emmanuel Raulin (Fr)

Joachim Weisbrod (Ge)

Ana Isabel Sanchez-Luengo (Sp)

Raquel del Rio Paramio (Sp)

Ole Black (Uk)

Mark Williams (Uk)

Steve Vale (Uk), temporarily OECD

John Perry (Uk)

Norbert Rainer (Au)

Zsolt Völfinger (Hu)

Márta Rónai (Hu)

Hans van Hooff (Ne)

Ton Bonné (Ne)