

# ВИЗИЯ ЗА ОРГАНИЗАЦИЯТА И ФУНКЦИОНИРАНЕТО НА НСИ ПРЕЗ 2025 ГОДИНА

*Деян Славов\**



## **1. Въведение**

Статистическата система е сложна система за събиране, обработка и анализ на статистическа информация. Тя се стреми да осигурява сравними, надеждни и отговарящи на строгите изисквания за конфиденциалност и защита на информацията данни, произведени в съответствие с международните стандарти. Ето защо е назряла необходимостта от модерна и използваща нови технологии статистика със съвременни средства за комуникация с респондентите, събиране на първична информация, нейната обработка и предоставянето ѝ на обществото. Целта на настоящия анализ е разглеждане на една визия за организацията и функционирането на НСИ след 2025 година. Описани са статистическите модели, използвани в ЕС, и тяхното развитие в бъдеще.

### **Някои факти и насоки за предизвикателствата, които предстоят в бъдеще:**

Според изследване на IDC, обхванало 11 вертикални индустрии, вкл. търговия на дребно (ритейл), здравеопазване, транспорт, правителствен сектор и други:

- Анализаторите очакват сегментът на автомобилите с разширена функционалност - т.нар. свързани коли (Connected Car) - да нарасне само в САЩ с 34% тази година.

---

\* Директор на ТСБ - Варна; e-mail: D.Slavov@nsi.bg.

- През 2013 г. пазарът на IoT се оценява на 1.3 трилиона долара.
- Очаква се в близките години този пазар да нараства с 13% годишно до 3.04 трилиона долара през 2020 г., а общият брой IoT устройства да достигне 30 млрд. броя.

**Според прогноза на Gartner:**

- Към 2020 г. към интернет ще бъдат свързани 26 милиарда всевъзможни устройства, докато през 2009 г. броят им е бил около 900 милиона.
- IoT е новата „гореща вълна“ в индустрията и големите ИТ компании започват да заемат позиции на пазара.

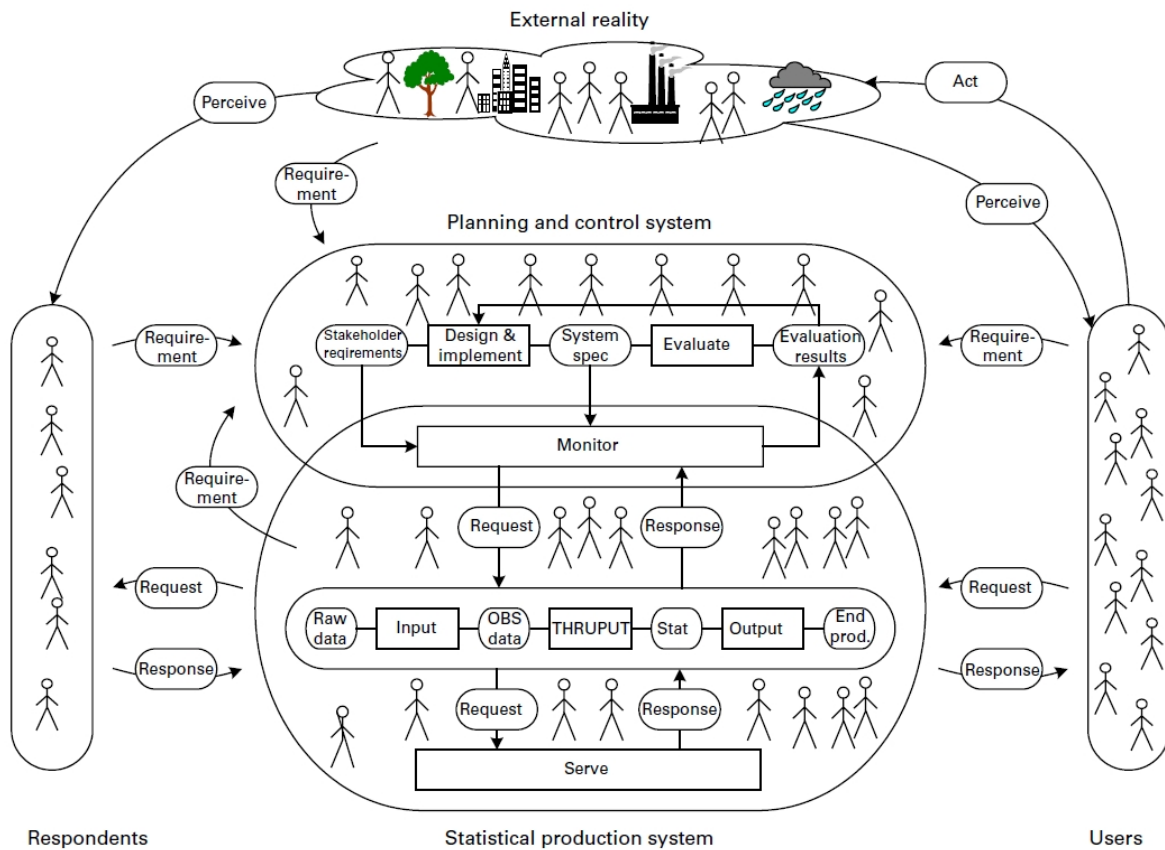
**Според учени от Харвардския университет:** Имплантирането на гъвкава електроника в мозъка е напълно възможно. Заедно с изследователи от Националния китайски център по нанонауки и технологии те са успели да инжектират електроника със спринцовка директно в тъканта на главния мозък. Създадената за целта технология представлява мрежа от гъвкави проводници и електронни елементи, които могат да се въвеждат директно в мозъка с помощта на спринцовка.

**Според инженерите на Intel:** Хората ще могат да управляват компютри чрез имплантирани в мозъка им чипове още през 2020 година.

**По данни на Мтел за 2014 година:**

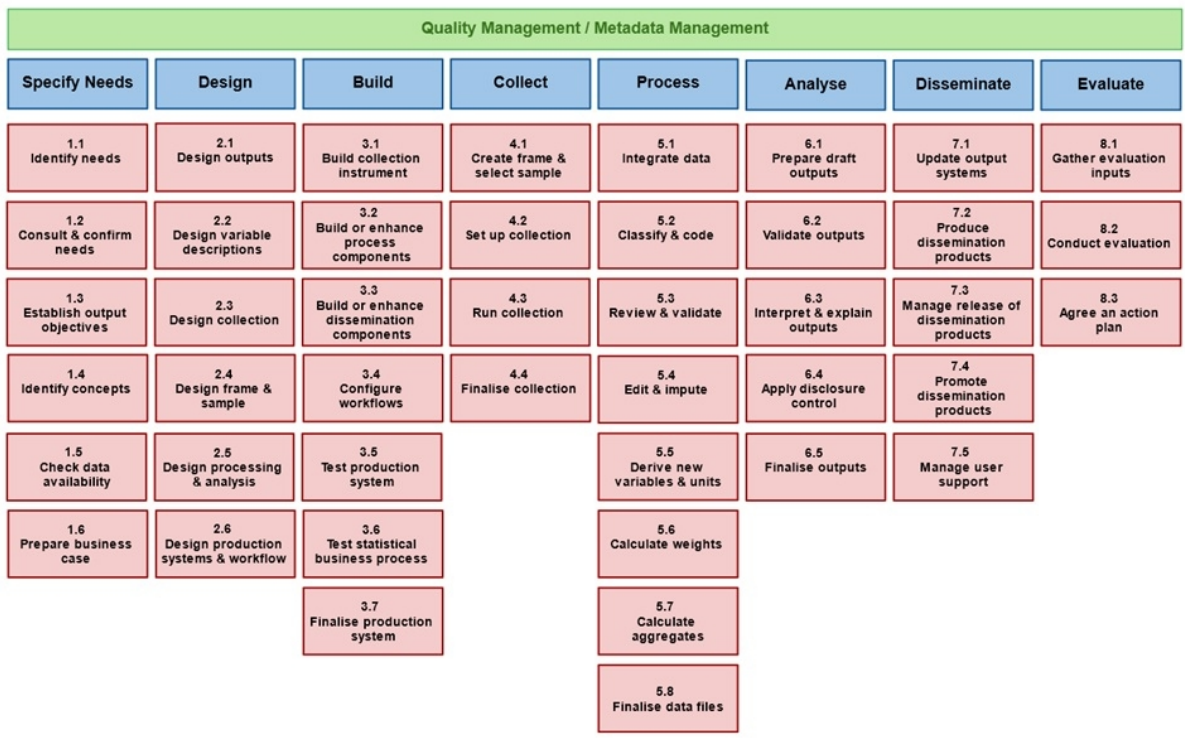
- 2 000 000 мобилни телефона в България са били смартфони.
- Смартфоните разполагат с възможност да предават данни за ползвателя като местоположение, честота на обажданията, използване на приложения и други с цел генериране на статистически данни без намесата на потребителя.
- При инсталация на безплатен софтуер от Google Play задължително се приемат условията на производителя на софтуера.
- Използване на платформата на фейсбук за програмиране на анкети.

Фигурата на Bo Sundgren в публикацията „Някои основи на статистическите системи“ от 2004 г. показва статистическата система като прецизирана структура, която измерва външната реалност чрез сложна система от взаимовръзки между служители, респонденти и потребители:



Процесите в системата за планиране и управление съответстват на фази 1 - 3 в GSBPM (Generic Statistical Business Process Model). Процесите в системата за производство на статистическа информация съответстват на фази 4 - 8 в GSBPM\*.

**\* Бизнес модел на статистическото производство**



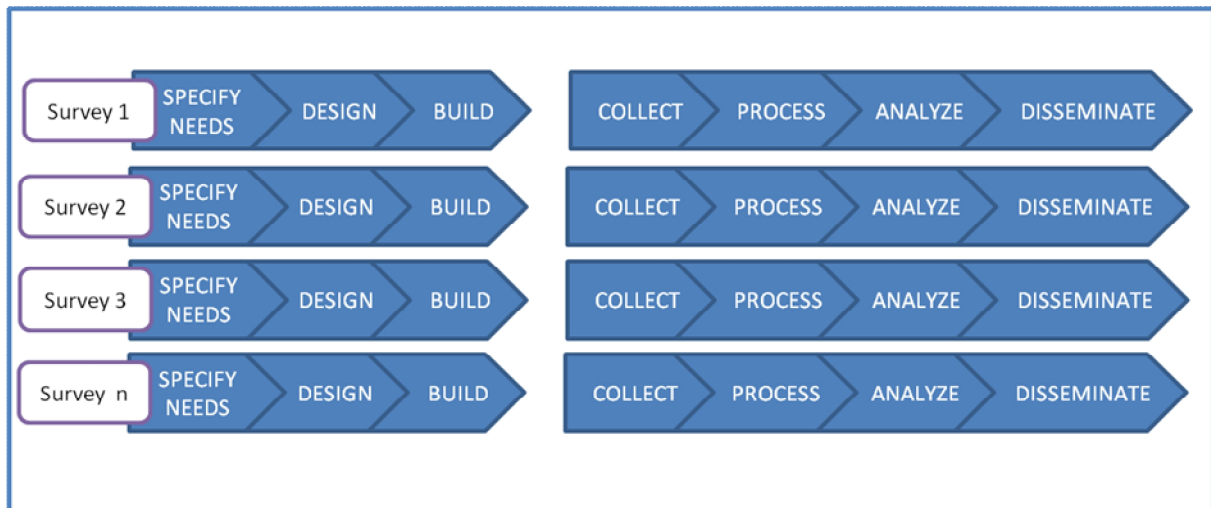
Тази схема визуализира най-общата стандартизирана структура на статистическата система. Чрез нея може да се представи организацията както на отделно статистическо изследване, така и на отделна статистическа служба - териториална или национална, и дори организацията на Евростат.

На практика съществуващите и функциониращи статистически системи се различават значително една от друга. Външните условия, в които те функционират, също се различават, като промените в тях стават все повече и все по-глобални. Това поражда необходимостта от промяна и адаптация и на статистическите системи независимо от това колко добре са работили до момента.

## 2. Статистически модели на производство

### 2.1. Тръбен модел

Най-разпространеният днес модел на производство в статистическите системи е тръбният модел (stovepipe). Той е резултат от един исторически процес, разработен самостоятелно във всички сектори. В тръбния модел работата по отделните статистически изследвания е независима като формат от други изследвания в почти всеки етап от веригата за производство на статистическа информация.



#### Предимства на тръбния модел

1. Производствените процеси са добре адаптирани към съответните продукти.
2. Той е гъвкав и може да се адаптира бързо към относително малки промени в явленията, чиито данни отчита.
3. Той е под контрола на ръководителя на изследването, което и предполага нисък риск при проблем в някой от производствените процеси и обикновено не следва да засяга останалата част от производството.

#### Недостатъци на тръбния модел

1. Може да доведе до ненужно натоварване на респондентите, когато събирането на данни се извършва несъгласувано и от респондентите се изисква една и съща информация повече от един път.
2. Тръбният модел не е добре пригоден да събира данни за явления, които се описват с множество различни величини като глобализацията, устойчивото развитие и изменението на климата.
3. Не на последно място този начин на производство е неефективен и скъп, тъй като той не се възползва от стандартизацията между регионите и сътрудничеството между държавите членки. Тръбният модел е доминиращият модел в Европейската

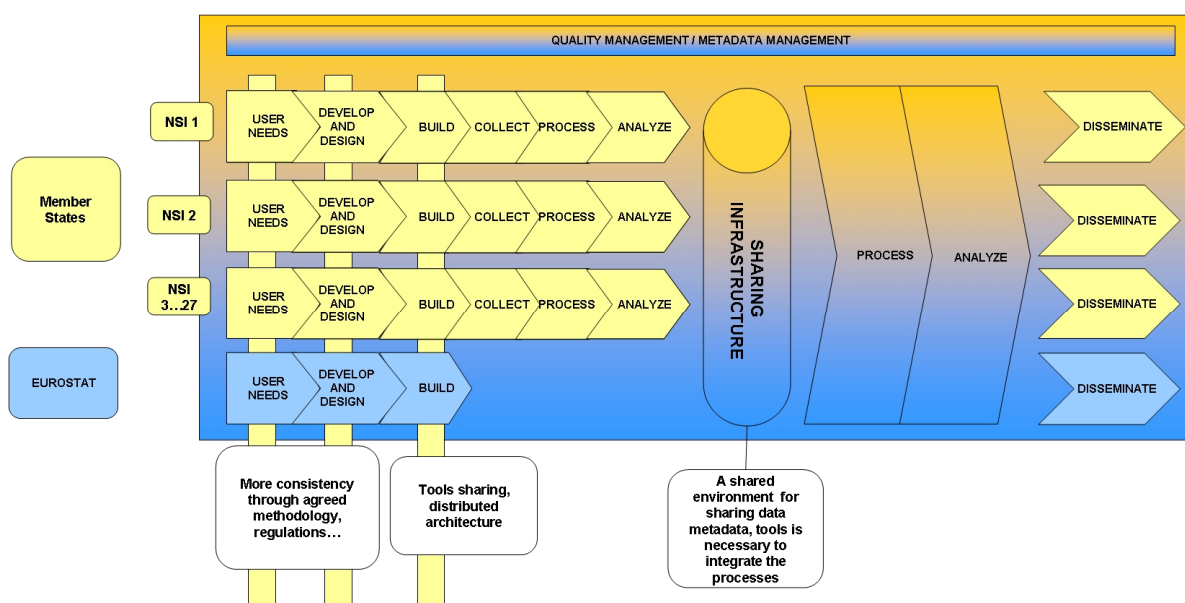
статистическа система (ЕСС) и е възпроизведен дори на равнище Евростат като уголемен тръбен модел.

## 2.2. Интегриран модел

Интегрираният модел е нов и иновативен начин за изготвяне на статистика. Той се основава на комбинация от различни източници на данни. Тази интеграция може да бъде хоризонтална или вертикална.

**1. Хоризонталната интеграция** е между статистически области на ниво национални статистически институти и Евростат. Хоризонтална интеграция означава, че европейската статистика вече не произвежда информация, събирайки данни за всяко статистическо изследване поотделно от даден отделен източник, а по интегриран начин, който съчетава индивидуалните особености на различни изследвания/източници в процеса на събиране на статистически данни на по-ранен етап например домакинства (W. Radermacher, A. Baigorri, D. Delcambre, W. Kloek, H. Linden, 2009).

**2. Вертикалната интеграция** е едновременно на национално и европейско равнище. Вертикалната интеграция трябва да се разбира като плавно и синхронизирано функциониране на информационните потоци на национално ниво и на ниво ЕСС независимо от източниците (респонденти или администрация) до крайния продукт (данни или метаданни). Вертикалната интеграция се състои от два елемента: съвместни структури, инструменти и процеси и т.нар. европейски подход към статистиката (W. Radermacher, A. Baigorri, D. Delcambre, W. Kloek, H. Linden, 2009).



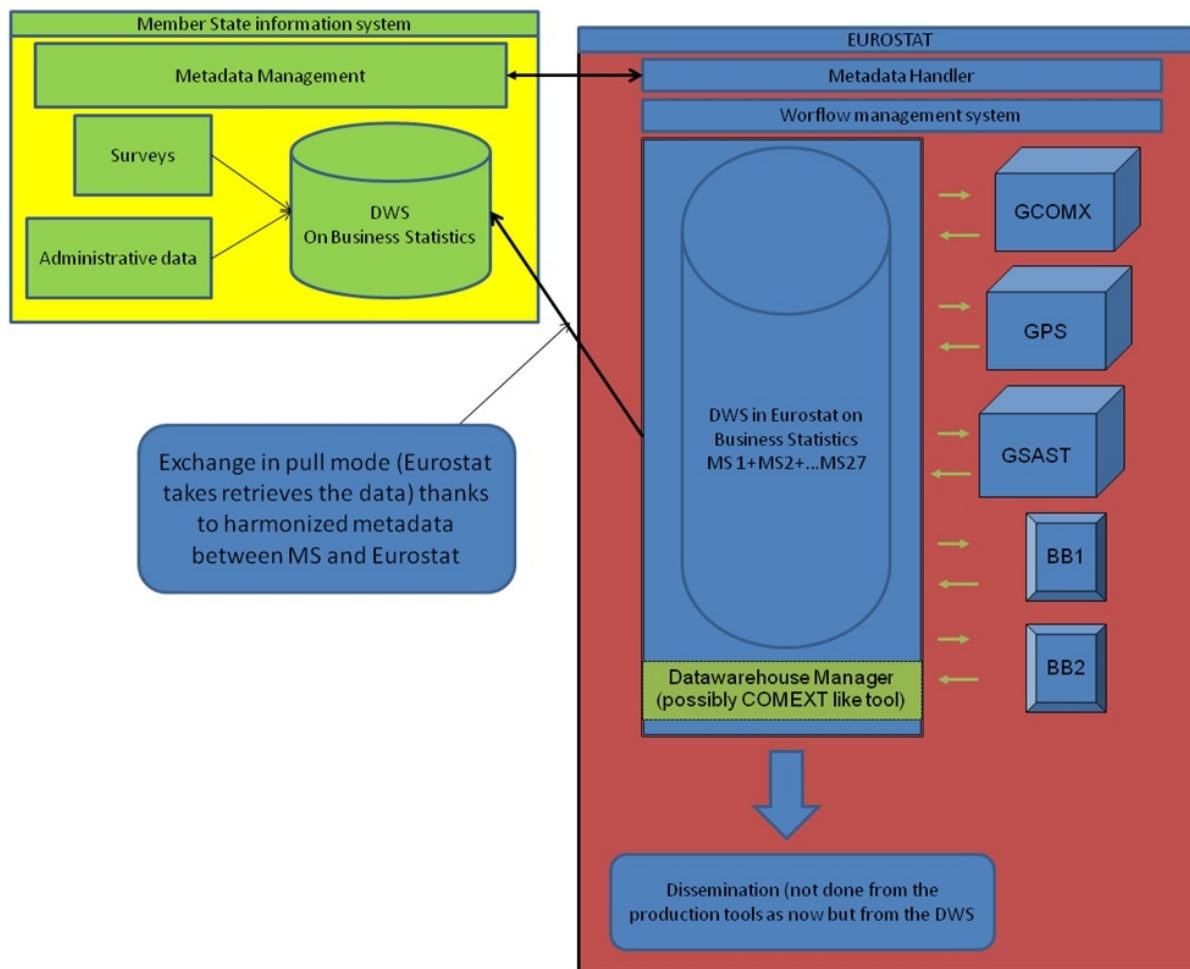
Интегрираният модел е създаден, за да се избегнат недостатъците на тръбния модел (натоварване на респондентите, неподходящи за проучване многомерни явления, неефективност и високи разходи). „Чрез интегрирането на набори от данни и комбиниране на данни от различни източници (включително административни източници) различните недостатъци на тръбния модел могат да бъдат избегнати. Този нов подход ще подобри ефективността чрез премахване на ненужната вариация и дублиране на дейности и създаване на свободни мощности за предстоящи информационни нужди” (W. Radermacher, A. Baigorri, D. Delcambre, W. Kloek, H. Linden, 2009).

А задачата за преминаване от тръбния към интегрирания модел никак не е лесна. В своя отговор до Статистическата комисия на ООН относно проекта „Насоки за интегрирани икономически статистики” В. Радермахер пише: „Преминаването от една принципно единна система като Системата от национални сметки към практически интегрирана система представлява дългосрочен проект и ще изисква интеграция в производството на първични статистически данни. Това е приоритетна цел за Евростат, декларирана чрез съобщение на Европейската статистическа система още през 2009 г. до Европейския парламент и Европейския съвет, за метода за изготвяне на статистическата информация на ЕС” („Визия за новото десетилетие”).

### **2.3. Warehouse подход**

В допълнение към тръбния модел разширеният тръбен модел и интегрираният модел W. Radermacher, A. Baigorri, D. Delcambre, W. Kloek, H. Linden (2009) описват също т.нар. подход **склад за данни** (datawarehouse): „Подходът „склад за данни” осигурява получаването и съхранението на данни веднъж, но тези данни се използват за различни статистически цели. Складът за данни третира информацията като актив за многократна употреба. В основата си моделът на склада за данни не е специфичен за определено отчетно или аналитично изследване. Вместо да се концентрира върху един процес, ориентиран към определено изследване, основното хранилище е моделирано въз основа на данни за вътрешни връзки, които са от съществено значение за организацията в целия процес”.

## Концептуален модел на склад за данни в Европейската статистическа система



Източник: W. Radermacher, A. Baigorri, D. Delcambre, W. Kloek, H. Linden (2009).

„Въз основа на такъв подход статистиките за конкретни области не трябва да бъдат произведени независимо една от друга, а като интегрирани части от цялостна производствена система, наречена „склад за данни“. Складът за данни може да се определи като централен регистър за данните, събрани чрез различни източници“ (W. Radermacher, A. Baigorri, D. Delcambre, W. Kloek, H. Linden (2009).

### 3. Интегриран warehouse модел

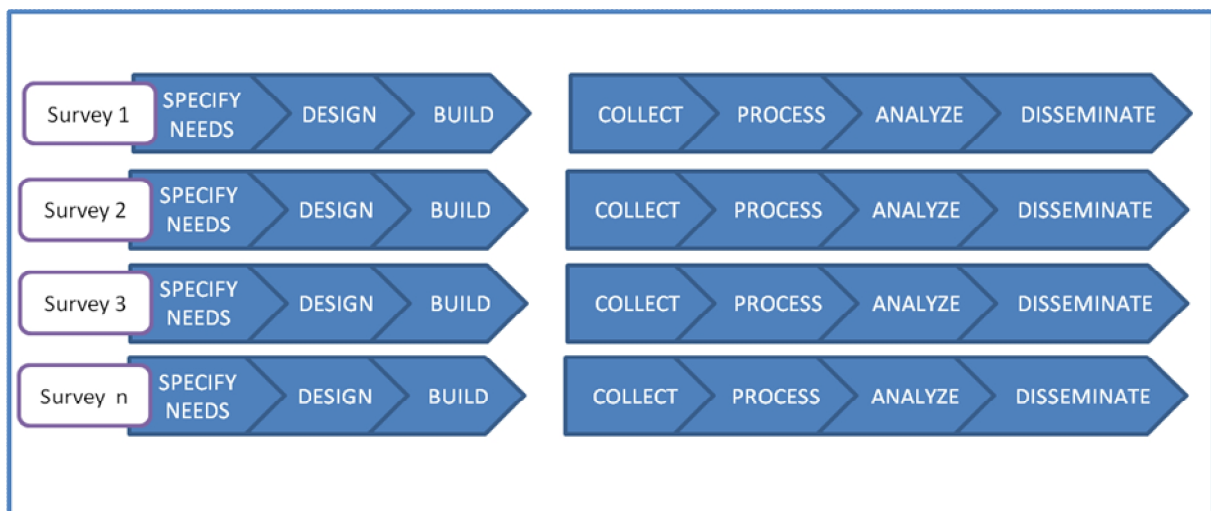
Моделът на интегрирания склад за данни (Integrated Warehouse) съчетава интегрирания модел и склада за данни в един модел. За да има интегриран склад, ориентиран към нуждите на статистическата система за производство на информация от различни области, трябва да има по-голяма съвместимост на методологията, да се



използват общи инструменти и достъпна архитектура. Първо ще разгледам интеграцията, после склада и след това как да се комбинират и двете в един модел.

### 3.1. Техническа интеграция на платформата

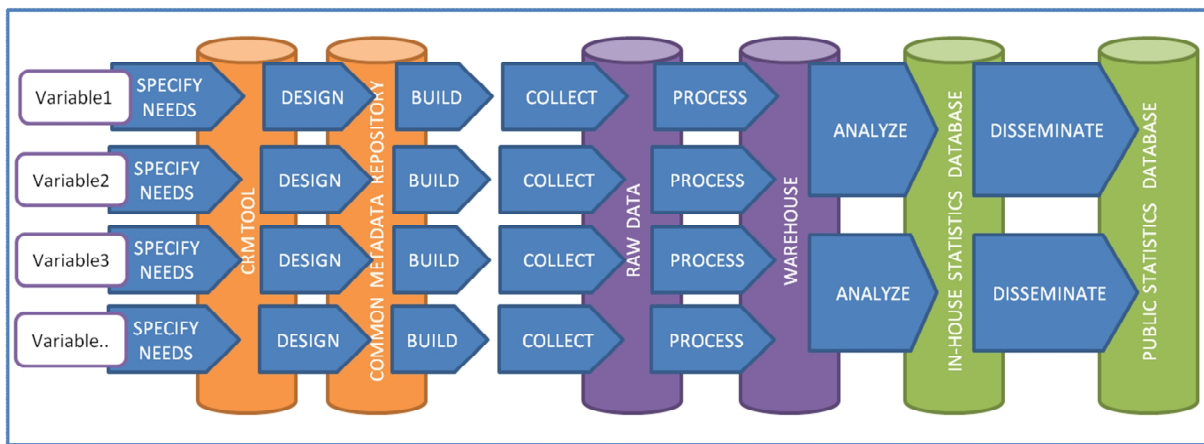
Нека да разгледаме класическата производствена система и да се опитаме да намерим ключови точки за интеграция, в която статистическите дейности си съответстват. Класическият тръбен модел на статистическа система изглежда по следния начин:



Нека започнем интеграцията на платформата от края на производствената система. Всяка такава единна система за разпространение на статистическа информация има за основа база данни, която включва всички подробни статистически данни. Едни са за вътрешно ползване, а други - за обществено ползване. За да се получи разнообразна и всеобхватна информация, особено от кръстосани изходи, ние се нуждаем от този склад, където данните се съхраняват веднъж, но могат да се използват за различни цели. Такъв склад трябва да бъде между фазите на обработка и анализ на данните.

В зависимост от конкретните използвани инструменти или други обстоятелства може да има повече от един склад от данни или база данни, предназначени за разпространение, но по-малко е за предпочитане. Например Статистическата служба на Естония има три интегрирани склада с бази данни. Първият е уеббазиран инструмент за събиране на данни от предприятията. Вторият е система за събиране на данни за социални изследвания. Третият е за административни и други източници на данни.

Но това не е всичко. Нека да разгледаме фазите на планиране и проектиране. Описанията на всички статистически действия, всички класификатори, които са в употреба, входните и изходните променливи, избрани източници на данни, описания на изходни таблици, въпросници и др., всички тези метаобекти по време на фазите на проектиране и изграждане следва да бъдат събрани в единно хранилище за метаданни. Заявките на потребителите трябва да се съхраняват в централната база данни - CRM (Customer Relationship Management) база данни.



Посочени са основните точки за интеграция в нивото на база данни. Но това не е нещо ново или революционно. Софтуерните инструменти могат да бъдат поделени между статистическите дейности. Възникват въпросите: От колко системи за събиране на данни се нуждаем? От колко системи за обработка или инструменти за разпространение се нуждаем? Имаме ли нужда от различен софтуер за обработка за всяко статистическо изследване или за всяка статистическа служба? За да се отговори на тези въпроси, трябва да се навлезе по-навътре в процесите и да се намерят начини да се стандартизират подпроцеси и методи. Един от начините за това е да се разгледат променливите във всяка статистическа дейност.

### 3.2. Процесна интеграция

„Интеграцията следва да обхваща всички етапи на производствения процес - проектирането на системата за събиране, обработката и разпространението на информацията” (W. Radermacher, 2011). За всяко статистическо изследване се проектират извадки и въпросници, използват се различни класификатори, определят се източниците на данни в зависимост от статистическото изследване.

В статистическата система има редица статистически дейности и всяка дейност събира някакви входни променливи и произвежда някакви изходни променливи. Един от начините да се намерят някои допирни точки между различните статистически дейности и източници е да се обърне внимание на променливите. Най-вече на входните променливи, защото събирането и обработката на данни са най-скъпите фази на производството на статистическа информация. Стандартизирането на тези фази ще доведе до най-бързата и най-голяма икономия. Изходните променливи могат да бъдат стандартизирани чрез инициативата SDMX (Statistical Data and Metadata Exchange).

Статистическите дейности следва да събират уникални входни променливи не само по редовете и колоните на таблиците във въпросника, а всяка променлива трябва да бъде събрана и обработена веднъж и запазена за съответния период. Това е необходимо, за да могат входящите променливи в склада за данни да бъдат използвани за производство на различни изходи. Този ориентиран към данните фокус предизвиква промени в почти всички фази на статистическия производствен процес. Извадките, анкетите, правилата за обработка, методите за импутация, източниците на данни и т.н. трябва да бъдат проектирани и произведени в съответствие със стандартизираните входни данни, а не според нуждите на едно конкретно статистическо действие или изследване.

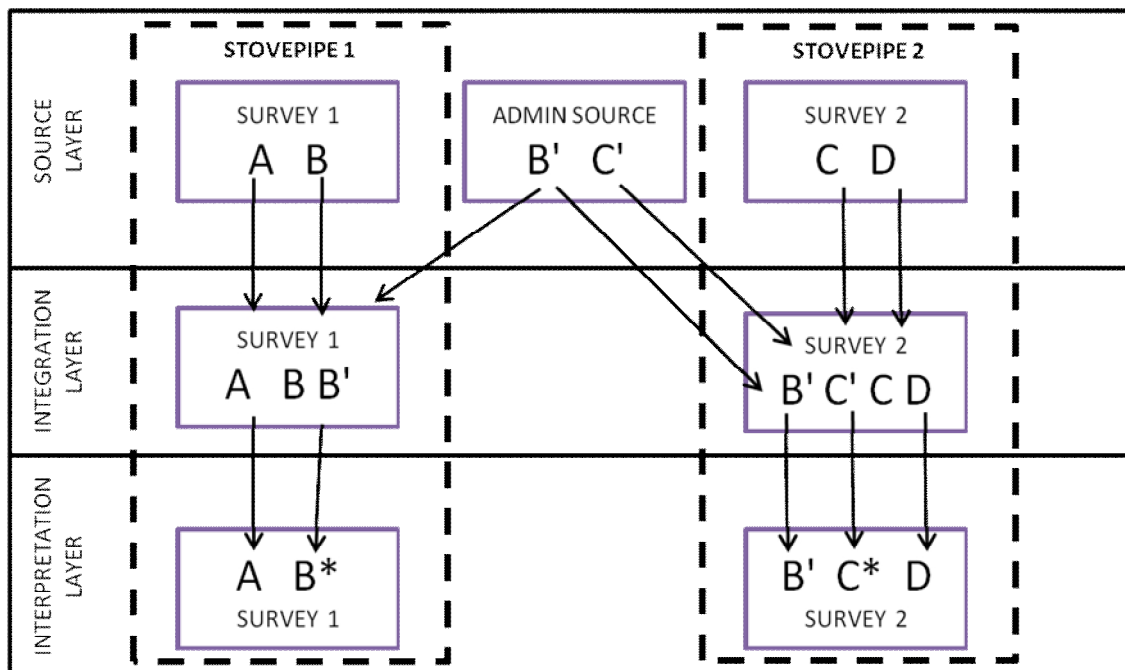
Променливите, събирани по такава система за производство на статистическа информация, ще доведат до намаляване на административната тежест, ще понижат разходите за тяхното събиране и обработка, ще дадат възможност да се произвежда по-разнообразна и по-навременна статистическа информация. Разбира се, това е ограничено в един стандартизиран дизайн. Ако има нужда от специално проучване, може да се проектира собствена извадка, въпросник и др., като това означава отделен проект със своя собствена цена. Но не е разумно да се произвежда по този начин регулярна статистическа информация, за която е необходимо всяко изследване да има собствен дизайн на извадка, въпросник и други.

### **3.3. Warehouse - повторно използване на данни**

За да се организира повторно използване на вече събрани и обработени данни чрез такава система за производство на статистическа информация, статистическите дейности не трябва да бъдат ограничени в рамките на нуждите на едно статистическо изследване. Статистическите дейности са свързани със събирането на входни данни, с методите за обработката им и получаването на изходни променливи. Когато говорим за

данни и повторното им използване, ние се интересуваме от събраните данни, извадките, с които е работено, отрязъка от време, за който са събрани тези данни.

Следващата фигура представлява типичен сценарий с две проучвания и един административен източник на данни. Проучване 1 събира с въпросници два вида входящи данни А и В и може да се използва информацията В' от административния източник. Проучване 1 анализира данни А и В\*, където В\* е информация В от въпросника или В' от административен източник. Проучване 2 събира данни С и D и анализира В', С\* и D.

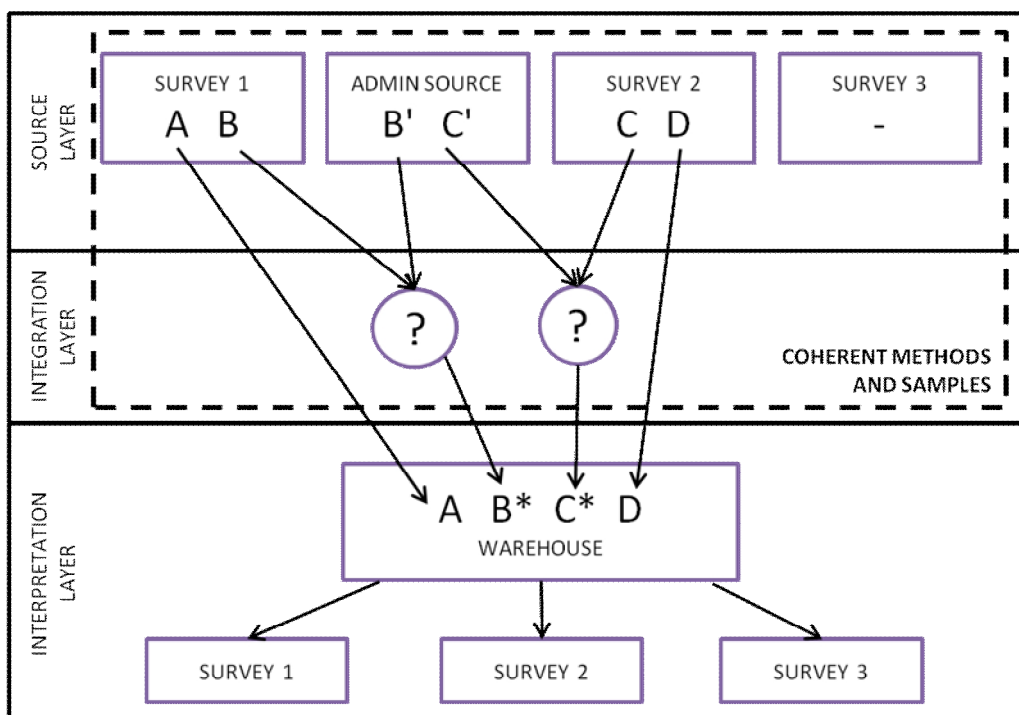


Това са статистическите действия въз основа на тръбния модел. В този случай е трудно да се използват повторно данните от слоя на интеграция, защото изчисляването на условните стойности в слоя на интеграция за В\* и С\* е направено за конкретно условие и няма голям избор от подобни променливи в слоя на интерпретация. Също така извадките от Проучване 1 и Проучване 2 може да не са обвързани, което означава, че трето изследване, което да стъпи на анализа на променливи А, В' и D в слоя на интерпретация, без да ги събира отново, ще има проблем при извадката и кохерентността на данните.

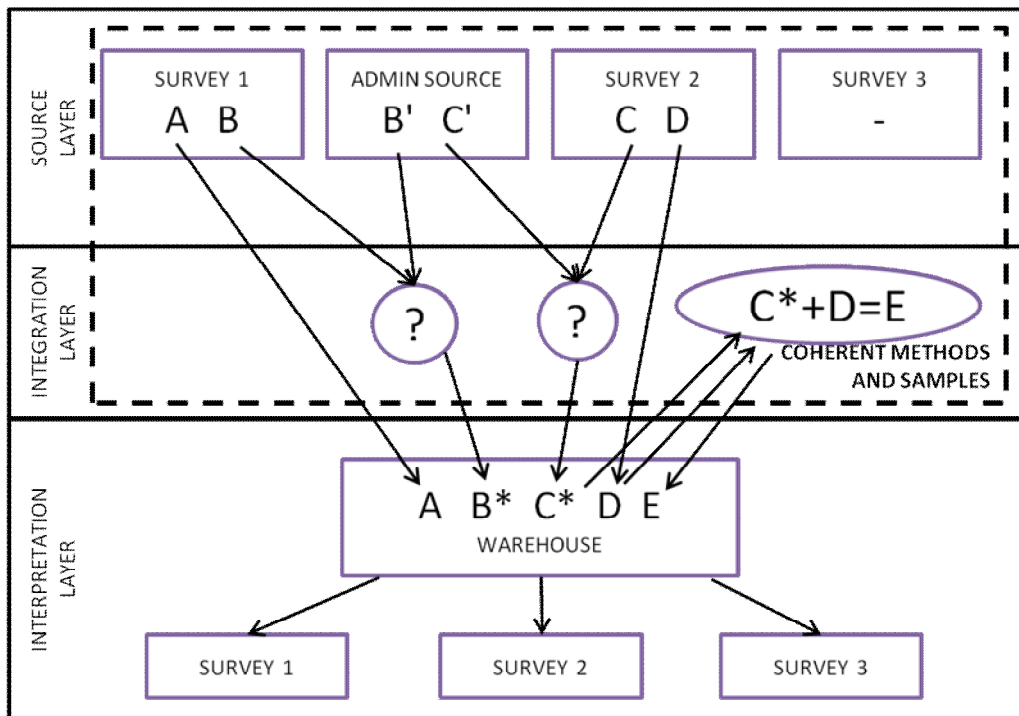
За решаване на проблема трябва да се инвестират време и усилия за планиране и подготовка на Проучване 1 и Проучване 2, така че те да бъдат съгласувани в уникална интегрирана извадково-информационна система - склад за данни (data warehouse).

В допълнение към анализа на данните и генерирането на изходни кубове с данни слой за интерпретация може да се използва и за достъп до данните на производство. В слой за интерпретация статистиците могат да планират и подготвят Проучване 1 и Проучване 2, като координират проучванията и архивите за обща рамка за оценка и определяне на уникални данни. Информацията, получена в този слой, е основа за разработване и настройка на редовни производствени процеси в слоя за интеграция. Това означава, че съгласуваният подход може да се използва, ако статистиците планират своите действия, следвайки логична йерархия на оценката на данните в обща рамка. Това, което ИТ специалистите трябва да поддържат впоследствие, е адекватна среда за проектиране на тази стратегия.

След това съобразно интегрирания подход Проучване 1 и Проучване 2, чрез които се събират данни от въпросници и един административен източник, могат да послужат за основа за последващи решения. Но този път решенията трябва да са на етап проектиране - проектиране на въпросника, избор на извадка, метод на импутация и др., като те се изготвят глобално, за всичките три проучвания. По този начин интеграцията на процесите ни дава данни, които могат да се използват многократно в склада за данни. Складът за данни съдържа всяка събрана информация само веднъж, което я прави много по-лесно използвана повторно.



Друг начин за повторно използване на данни, които вече са в склада, е да се изчислят нови показатели. Следващата фигура илюстрира сценария, при който новият показател Е се изчислява от показатели С\* и D, които са заредени вече в склада.



Това означава, че данните могат да се преместват назад от склада за данни към слоя на интеграция. Данните от склада за данни (Data Warehouse) могат да бъдат използвани в процеса на интеграция за множество цели, вкл. и за изчисляването на нови показатели.

Интегрираният склад за данни отваря пътя за нови възможности - статистически дейности, при които не трябва да се събират и обработват данни, а може да се произвежда статистическа информация директно от данните в склада за данни. Пропускането на етапите на събиране и обработка на данните води до много по-бързо производство на статистическа информация и анализи и не на последно място е много по-евтино отколкото в случаите на класическо проучване.

За проектирането и изграждането на система за производство на статистическа информация в съответствие с модела на интегрирания склад за данни първоначално се изискват повече време и усилия отколкото при изграждането на тръбния модел. Но разходите за поддръжка на интегрирания склад за данни са по-малки, а статистическите

продукти се произвеждат по-бързо и по-евтино, което компенсира първоначалните инвестиции.

#### **4. Статистическият склад за данни (S-DWH) като слоеста модулна система**

##### **4.1. Слоеста (Layered) архитектура**

В същността на статистическия склад за данни (S-DWH) идентифицираме четири функционални слоя, в които са групирани дейностите. Входният слой съответства на процесите, при които от външни източници постъпва информация по съответните интерфейси, а в изходния слой произведените данни са публикувани за външни потребители. В междинните слоеве на системата можем да идентифицираме дейностите по извличане, трансформиране и зареждане (ETL) в склада за данни (DWH), извършване на анализ, извличане на данни, дизайн за възможни нови стратегии или данни за повторна употреба. По-конкретно, от входа до изхода на склада за данни разграничаваме:

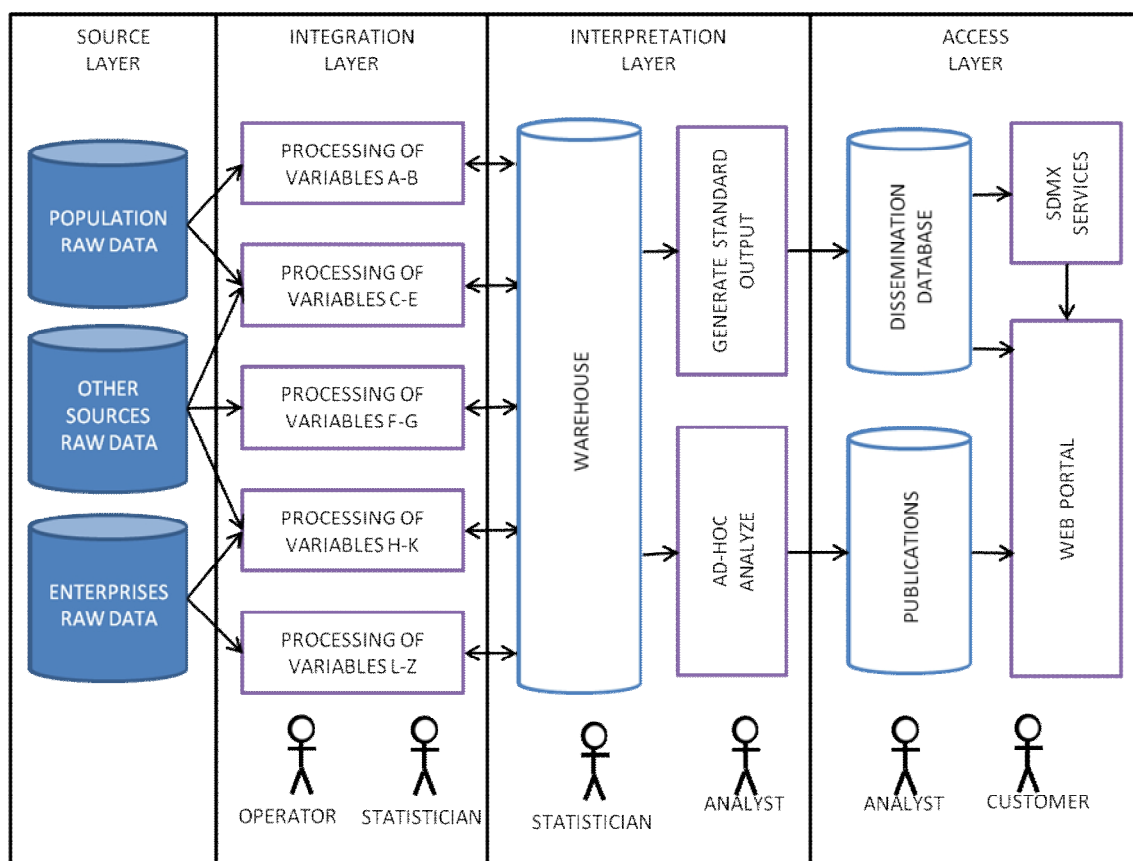
**Слой IV - слой за достъп** на окончателното представяне на данните, разпространение и предоставяне на исканата информация.

**Слой III - слой за анализ и тълкуване на данните;** той е специално за статистики и позволява всякакъв анализ на данни, извличане на данни и описване на бизнес процесите или данните за повторна употреба.

**Слой II - слой за интеграция,** където се извършват всички оперативни дейности, необходими за всеки статистически производствен процес.

**Слой I - слой за източници на данни,** нивото, в което ще поставим всички дейности, свързани със съхраняване и управление на вътрешни или външни източници на данни.

Слоеве в статистическия склад за данни (S-DWH) са в определен ред и данните преминават през всеки слой, без да прескачат слоеве. Не е възможно да се използват данните директно от другия слой. Ако данните са необходими, те трябва да бъдат преместени в слоя, където са необходими. И те не могат да се преместват така, че някои слоеве да се пропускат. Данните могат да бъдат премествани само между съседни слоеве.



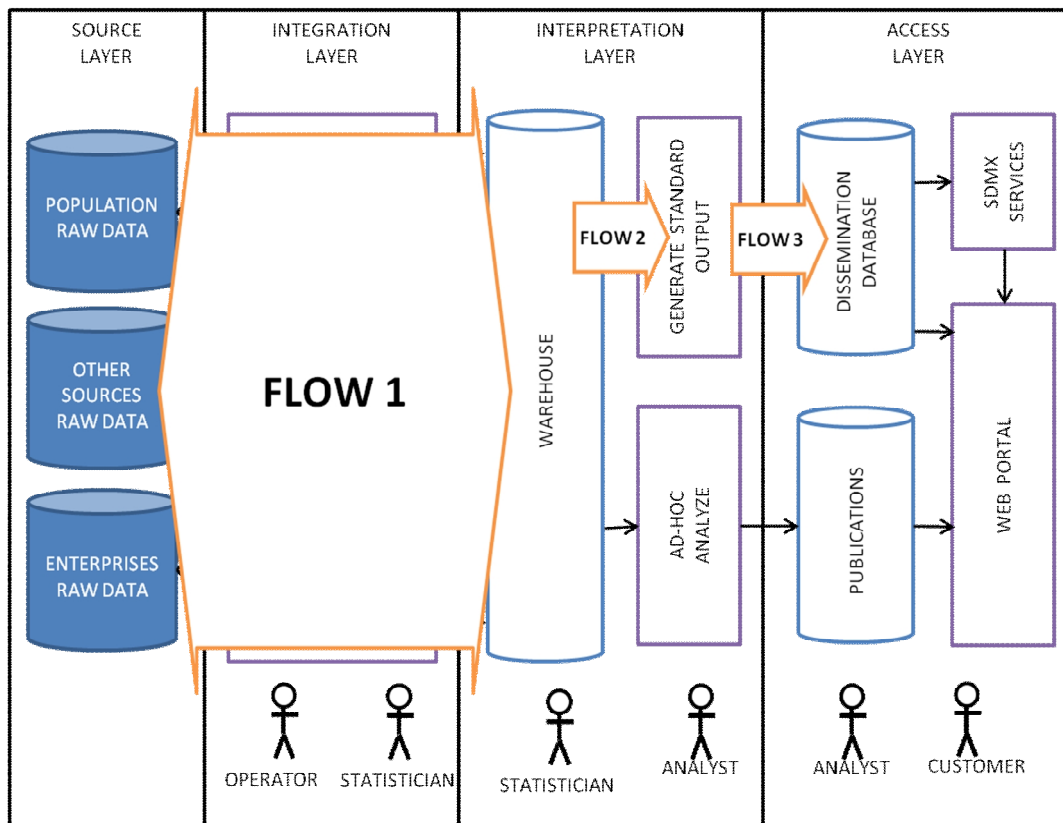
Например за да се публикуват данни в слоя за достъп (access layer), първо трябва да бъдат събрани данни в суров вид в слоя за източници на данни (source layer), след което да бъдат заредени за обработка в слоя за интеграция, едва след това данните се зареждат в склада за данни (DWH), в слоя за анализ и тълкуване на данните (Interpretation layer) и чак тогава могат да се изчисляват статистически данни или да се направи анализ и да се публикува в слоя за достъп (access layer).

Друг пример: понякога е необходимо да се наблюдава процесът на събиране и анализиране на първичните данни по време на събирането. Тогава първичните данни се извличат от базата данни за събиране и се обработват в слоя за интеграция, така че данните да могат лесно да бъдат анализирани със специфични инструменти, заредени в слоя за интерпретация (interpretation layer), където могат свободно да бъдат анализирани. Този процес се повтаря толкова често, колкото е необходимо - например веднъж на ден, веднъж седмично или почасово.



#### 4.2. CORE (COmmon Reference Environment) услуги и повторно използване на компоненти

Има три основни групи от работни потоци в статистическия склад за данни (S-DWH). Единият поток актуализира данните в склада за данни, вторият актуализира базата данни за разпространение за стандартни справки (данни), а третият поток актуализира базата данни за публично разпространение.



Тези три автоматизирани работни потока от данни са напълно независими един от друг. Първият поток е най-големият и най-сложният. При него се извличат първичните (сурови) данни от слоя за източници на данни (source layer), обработват се в слоя за интеграция (integration layer) и се зареждат в слоя за анализ и тълкуване (interpretation layer). От друга страна, първият поток (Flow 1) предоставя чисти данни на слоя за източници на данни (source layer) за предварително попълване на въпросници, подготвя извадки за системите за събиране и т.н. Нека да наречем този поток „поток за преработка“.

Вторият и третият поток са много сходни, те генерират стандартен изход за бази данни за разпространение. Единият актуализира данните във „вътрешна“ база от данни

за разпространение, а вторият - в публична база данни. И двата потока са еднопосочни. Нека наречем втория поток „куб за генериране на данни“, а третия поток - „куб за публикуване на данни“. В този смисъл кубът е многомерна таблица подобно на таблица от софтуерите Stat или PC-Axis (<https://www.h6.scb.se/pxwebtest/pxweb/en/>).

Потоците трябва да бъдат изградени около входните данни или групи от входни данни, за да могат да захранват склада за данни. Кубовете за генериране на данни и публикуване на данни са изградени така, че всеки поток генерира данни в куб или публикува данни в куб.

Има много софтуерни инструменти за изграждане на тези модулни потоци. Самата слоеста архитектура на статистическия склад за данни (S-DWH) предоставя възможност за използване на различни платформи и софтуер в отделните слоеве, т.е. да се използват компоненти, с които НСИ вече разполага, или такива, прилагани в други национални статистически служби. В допълнение могат да се използват различни софтуерни продукти в един и същ слой, за да се изгради един конкретен поток. Проблемите възникват, когато се опитваме да използваме тези различни модули и различни формати на данни заедно.

Обикновено CORE (COmmon Reference Environment) е среда, определяща статистическите процеси и автоматизираща тяхното изпълнение. Възможно е инструмент, разработен от друга национална статистическа служба по принципите на CORE, лесно да се интегрира в рамките на статистическия процес на друга статистическа служба. Освен това наличието на единна среда за изпълнение на всички статистически процеси осигурява високо ниво на автоматизация и пълна възпроизводимост на процесите на изпълнение.

Всички статистически служби произвеждат официална статистическа информация, споделят сходни цели, а оттам и извършват едни и същи дейности, свързани с изготвянето на статистическа информация. Независимо от това дейностите в момента се извършват самостоятелно, без да се разчита на съвместни решения. Приемането на обща архитектура ще доведе до намаляване на разходите, дължащи се на дублирани дейности, както и до подобряване на качеството на произведените статистически данни поради приемането на стандартизирани решения.

### **Основните принципи, залегнали в CORE дизайна, са:**

**а) Независима платформа.** Статистическите служби използват различни платформи (например хардуер, операционни системи, системи за управление на бази данни, статистически софтуер и т.н.) и оттам архитектурата е обречена на неуспех, ако тя се стреми да наложи стандарти на техническо равнище. Освен това независимостта на платформата позволява да моделираме статистическите процеси на концептуално ниво, така че те не трябва да бъдат променяни, когато променяме услуги.

**б) Ориентирана към услуги.** Визията е, че изготвянето на статистически данни се осъществява чрез услуги, извикващи други услуги. Следователно услугите са градивните елементи на архитектурата. Като има ясна комуникация на интерфейсите (връзките) и в услугите се приложат принципите на съвременното софтуерно инженерство, се получава архитектура, ориентирана към услуги или SAAS.

**в) Слоест подход.** Според този принцип някои от услугите са вследствие на други дейности и са разположени в горната част на статистическия процес. Така например публикуването на данни е услуга, която изисква производството на всички видове данни, разположени по-рано в статистическия процес, като събиране на данни и съхранение на информация. Амбицията на този модел е да се свърже целият набор от слоеве - от събирането до публикуването, чрез описване на всички слоеве от гледна точка на услугите, предоставяни на по-високо ниво, по такъв начин, че всеки слой зависи само от първия долен слой.

Основната цел на CORE е разработването и прилагането на среда в подкрепа на определени статистически процеси и автоматизиране на тяхното изпълнение. Основните процеси са проектирани по стандартния начин, като се започне от предлаганите услуги. Всеки конкретен процес е осигурен по отношение на абстрактни статистически услуги, които могат да бъдат нанесени на специфични ИТ инструменти. CORE допринася за насърчаване на обмена на средства сред различните национални статистически служби. В действителност инструмент, разработен от една статистическа служба, следва да бъде в съответствие с основните принципи в производството на статистическа информация и по този начин лесно да бъде интегриран в рамките на статистическия процес в друга национална статистическа служба. Освен това наличието на единна среда за изпълнение на всички статистически процеси осигурява високо ниво на автоматизация и пълна възпроизводимост на процесите на изпълнение.

За нас е много важно да се направят някои преходи и съпоставяне между различни модели и подходи. За съжаление, схематичното представяне на CORE

процеса като бизнес модел не е възможно, защото моделът на CORE е информационен модел и няма начин да бъде изразен директно чрез бизнес модел. Двата модела са за различни неща. Те могат да бъдат свързани само ако тази връзка е по някакъв начин част от моделите.

Информационният модел на CORE е проектиран като начин на мислене. В рамките на този модел статистическата служба е обект, както и всяка една от нейните характеристики е препратка към GSBPM (Generic Statistical Business Process Model). Обикновено различните служби използват различни услуги със собствени инструменти, като получават различни формати за данни. Така че за да си взаимодействат услугите, трябва да се преобразуват. Очевидно е, че преобразуванията са скъпи. В най-общ смисъл решенията за интеграция на програмни интерфейси може да направи CORE услугата изпълнима. Основната услуга е наистина специфична и изходът за интеграция е чрез въвеждане и интеграция на изходни програмни интерфейси. Такива програмни интерфейси трансформират от/до CORE модела в конкретния формат на инструмента.

CORE преобразуванията са създадени за категории инструменти, а оттам и интеграционните програмни интерфейси би трябвало да поддържат желаните трансформации (например от CSV към CORE и от CORE към CSV). По принцип, приложният програмен интерфейс за интеграция се състои от набор от компоненти за трансформация. Всеки компонент за преобразуване съответства на определен формат на данните и основните елементи при проектирането им са специфичните файлове и трансформиращите операции.

С цел осигуряване на входни данни за единен инструмент се налага операцията по трансформация към CORE. От друга страна, на изхода има отново операция по трансформация към CORE. За всеки входен или изходен файл трябва да съществува възможност за трансформация. По този начин повторно използване на компоненти може да се извърши по много лесен и ефективен начин.

## **5. Заключение**

Днес преобладаващият модел за създаване на статистиката е тръбният модел. Но съществуват и интегрираният модел, и Warehouse подходът. Формата на интеграция може да се разглежда в три аспекта:

1. Техническа интеграция - интегриране на ИТ платформи и софтуерни инструменти.

2. Интеграция на процеси - интегриране на статистическите процеси като дизайн на проучването, извадки, обработка на данни и т.н.

3. Интеграция на данни - данните се съхраняват веднъж, но се използват за различни цели.

Когато съберем всички тези три аспекта на интеграция заедно, ще получим статистически склад за данни, който е построен на базата на интегрирана технология, използва интегрирани процеси за изготвянето на статистическа информация и ефективно използва повторно данни. И на тази база статистическият склад за данни може да се използва за управление и промяна на статистическия производствен процес чрез слоя за интерпретация на едни и същи данни. Това ще допринесе за подкрепата на нови стратегии за статистическа обработка и до прецизиране на статистическия склад за данни.

## **ЦИТИРАНА ЛИТЕРАТУРА:**

**ESSnet CORE** (COmmon Reference Environment), <http://www.cross-portal.eu/content/core-0>.

**European Union, Communication** from the Commission to the European Parliament and the Council on the production method of EU statistics: a vision for the next decade, COM (2009) 404 final,

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0404:FIN:EN:PDF>.

**Radermacher, W., A. Baigorri, D. Delcambre, W. Kloek, H. Linden** (2009). Terminology Relating To The Implementation Of The Vision On The Production Method Of EU Statistics, [http://ec.europa.eu/eurostat/ramon/coded\\_files/TERMS-IN-STATISTICS\\_version\\_4-0.pdf](http://ec.europa.eu/eurostat/ramon/coded_files/TERMS-IN-STATISTICS_version_4-0.pdf).

**Radermacher, W.** (2011). Global consultation on the draft Guidelines on Integrated Economic Statistics.

**Sundgren, B.** (2010). The Systems Approach to Official Statistics, Official Statistics in Honour of Daniel Thorburn, pp. 225 - 260, <https://sites.google.com/site/bosundgren/my-life/Thorburnbokkap18Sundgren.pdf?attredirects=0>.

**The Sponsorship on Standardisation** (2013). Standardisation in the European Statistical System.

**UNSC** (2012). Guidelines on Integrated Economic Statistics, <http://unstats.un.org/unsd/statcom/doc12/RD-IntegratedEcoStats.pdf>.