

**Источники первичных данных  
для официальной статистики.  
Смена приоритетов.**

А.Суринов, НИУ ВШЭ.



# описание экономического производства с учетом природного и социального факторов





Айсберг. Мы не видим его нижнюю часть, но знаем, что она есть и с помощью приборов и вычислений можем рассчитать ее размеры и форму.





## **Источники, составляющие основу современных статистических наблюдений:**

- статистическая отчетность;
- опросы респондентов интервьюерами;
- сбор данных регистраторами без опроса респондентов;
- бухгалтерская, налоговая и иная ведомственная отчетность;
- государственные регистры, реестры и иные административные записи;
- архивные данные и результаты статистических наблюдений, проведенных в прошлом;
- международные и зарубежные национальные базы данных и статистические публикации;
- большие данные.



# Большие данные – новый источник данных об обществе

БД - это огромные массивы многообразной неопределенно структурированной и неструктурированной информации, образующейся как побочный продукт коммерческой и некоммерческой деятельности, спонтанно возникающие и прекращающие свое существование.

Явление БД связано с появлением технологических возможностей собирать, хранить, передавать, преобразовывать и анализировать огромные массивы данных

В качестве главных характеристик БД выделяют «пять V»: объём информации, как количество записей и их характеристик (*volume*), скорость прироста и обработки (*velocity*), многообразие источников, носителей и форматов данных (*variety*), достоверности (*veracity*) и ценности (*value*).



Статистическая Комиссия ООН следующим образом классифицировала **источники возникновения БД**:

- источники данных, связанные с осуществлением негосударственной программы, например, в медицине или страховании;
- коммерческие источники данных по операциям с кредитными карточками и онлайн-операции;
- источники данных, связанные с работой сенсорных сетей, например, данные от изображений, полученных со спутников, данные с автодорожных датчиков и метеорологические данные от измерительных устройств;
- источники данных, связанные с работой регистрирующих устройств, например, сети мобильной телефонной связи и GPS;
- источники данных, связанные с поведением пользователей Интернета и социальных сетей и пр.



# Характеристики типов данных, пригодных для использования в официальной статистике

Характеристики данных	Статистические	Административные	БД
Тип данных	Структурированные	Структуриро-ванные, частично или полностью	Неструктури-рованные
Возможность доступа	Регламентируется	Регламентируется	Не регламен-тируется
Наличие утвержденных правил при создании и распространении	Нормативно-правовой акт	Нормативно-правовой или ведомственный акт	Решение организации-владельца/создателя
Скорость создания	Наличие временного лага после события	Наличие временного лага после события	В реальном времени или с небольшим временным лагом
Обеспечение качества (достоверность, прозрачность методологии, своевременность)	Национальная статистическая система	Ведомство	Нет обязательства
Наличие утвержденной методологии	Обязательно	Желательно	Нет обязательства
Устойчивость во времени, смена методов сбора и обработки требует обеспечения сопряженности динамических рядов	Обязательна	Необязательна, зависит от изменения нормативной базы	Отсутствует



## **Современные требования к данным, собираемым в целях статистики:**

- оценка событий на качественном уровне;
- включение в статистические оценки сведений о субъективном восприятии респондентом событий;
- включение респондента в процессы регистрации событий и их измерения;
- использование прогнозных оценок развития изучаемых событий;
- расширение практики сбора сведений о событии с разных точек зрения (метод «зеркальной» статистики);
- повышение оперативности оценок;
- расширение набора критериев для классификации событий и единиц статистического наблюдения за счет использования многомерных классификаций и объединения сведений из разных наблюдений;
- Сбор сведений, о внутренней структуре единицы наблюдения, и получение характеристик составляющих ее элементов;
- сбор сведений, позволяющих проводить мэтчинг данных, полученных из разных источников.



**Детализация** результатов статистических наблюдений (продукт, отрасль, вид деятельности, социальная, возрастная и пр. группы населения, территориальный разрез и др.) требует решение проблемы «малых областей статистического изучения» (Small Statistical Areas). Суть этой проблемы связана с ограниченными возможностями выборочных статистических наблюдений обеспечить достаточную репрезентативность оценок по малым частям обследованной совокупности, и сохранить конфиденциальность информации из-за высокой вероятности прямого или косвенного раскрытия респондента. Это требование стало особенно актуальным в связи с глобализацией рынков. Но именно сегодня в условиях цифровизации и сбора данных от респондентов в электронном виде, создались условия для формирования официальной статистики с высокой степенью дезагрегации. Электронный сбор первичных сведений не связан с ограничением позиций в форме статистической отчетности. Это позволяет получить сведения о производстве, запасах, продажах по дезагрегированной номенклатуре.

У пользователей сложился устойчивый спрос на **пространственную статистику**, предоставляющую статистические сведения, описывающие различные уровни географической классификации экономических явлений. Дифференциация экономического развития территорий внутри стран, различия в условиях жизни, требования бюджетного законодательства формируют спрос на такие данные у правительственных организаций и бизнеса.

Многие страны решают эту проблему формируя статистику на муниципальном уровне.

В некоторых странах разрабатывается информация по статистическим территориальным объединениям, например, Statistical Metropolitan Areas и Statistical Micropolitan Areas в США, или по регионам, сформированным в соответствии с Nomenclature Territorial Units for Statistical Purposes (NUTS) в ЕС. Имеется опыт официальной статистики по городам.

## **Профилирование единиц статистического наблюдения**

В практику мониторинга и анализа социально-экономического развития внедряются новые показатели и классификации.

Стали востребованы комплексные характеристики единицы наблюдения (предприятие, заведение, домохозяйство, семья), отдельные признаки которой получены из разных источников и представляют собой агрегированные ее характеристики или их сочетания.

Это дает возможность описать единое предприятие как совокупность местных единиц или заведений. А реализация методологии демографии предприятий или оценки их инновационной активности усложняет описание предприятия за счет подключения дополнительных критериев.

Похожая ситуация и в области статистики домохозяйств, когда требуется описать домохозяйство исходя из характеристик отдельных лиц, в нем проживающих, например, при идентификации бедного домохозяйства через наличие лишений у отдельных его членов (детей), а не только у домохозяйства в целом.

В официальной статистике довольно давно использует метод прикладной социологии, когда выясняются мнения респондентов относительно того или иного социально-экономического явления. наиболее ярким примером служат **обследований деловой активности и обследования потребительских ожиданий намерений).**

Основными задачами обследований является:  
изучение мнения бизнеса и потребителей в лице населения о динамике общей экономической ситуации, проблемах , препятствующих экономическому развитию, материальном положении.

При опросе выясняются мнения производителей, торговцев и потребителей не только об их оценках прошлого и настоящего, но и о перспективах. Результаты обследования используется для прогнозирования на основе расчета опережающих индикаторов, отражающих совокупное мнение бизнеса и обобщенную оценку потребителей будущего состояния экономики.

Ответы респондентов представляют собой оценки по пятибальной шкале, в виде :рост(лучше), скорее рост, без изменения, скорее снижение, снижение (хуже).

То, что **опросы населения** являются одним из главных источников данных для статистики признали еще в 1853г. на первом Международном статистическом конгрессе (сегодня Международный статистический институт - МСИ), посвященном методическим вопросам организации обследований бюджетов домохозяйств рабочих, где были приняты рекомендации по программе подобных обследований.

Многие положения этих рекомендаций не устарели и реализуются сегодня.

Метод сбора – личный опрос респондента.

Респондент является представителем домашнего хозяйства, а домашнее хозяйство - единицей статистической совокупности.

Респондентская сеть должна состоять из типичных домашних хозяйств и/или соответствовать генеральной совокупности.





## Пример участия респондента в регистрации событий – **обследование использования суточного фонда времени (бюджета времени)**

Бюджет времени дает возможность количественных измерений затрат времени населением на различные виды деятельности, связанные с оплачиваемой занятостью, учебой, проведением домашних работ и свободного времени. Показатели использования суточного фонда времени бюджета времени используются для оценки гендерного равенства, межпоколенных отношений в семье и денежной затрат времени на неоплачиваемую занятость.

Респонденты фиксируют свои затраты времени отдельно в течение буднего дня и дня и в течение выходного (праздничного) дня. Для обеспечения международной сопоставимости результатов при агрегации затрат времени используются гармонизированные классификации видов деятельности.

Такие обследования проводят страны ЕС, а также рекомендуется проводить всем странам-членам ОЭСР на регулярной основе.

# Руководство по измерению доступа к ИКТ и их использования на уровне домашних хозяйств и отдельных лиц Издание 2020 года



Ниже представлены типовые вопросы для получения показателей, рекомендуемых МСЭ.





+

<b>Номер показателя</b>	<b>Формулировка типового вопроса, связанного с соответствующим показателем</b>
НН1	Имеет ли данное домохозяйство радиоприемник?
НН2	Имеет ли данное домохозяйство телевизор
НН3	Имеет ли данное домохозяйство фиксированную телефонную линию? Имеет ли данное домохозяйство мобильный телефон? Имеет ли данное домохозяйство смартфон?
НН 4	Имеет ли данное домохозяйство компьютер (настольный, портативный, планшетный или аналогичный)?
НН 5	Пользовались ли вы компьютером (настольным, портативным, планшетным или аналогичным) в любом месте в течение последних трех месяцев?
НН 6	Имеет ли данное домохозяйство доступ к сети Интернет?
НН 7	Пользовались ли вы Интернетом в любом месте в течение последних трех месяцев?
НН 8	Где Вы пользовались Интернетом в последние три месяца?
НН 9	Для какого каких видов деятельности Вы использовали Интернет в частных целях (в любом месте) за последние три месяца?
НН 10	Пользовались ли Вы мобильным телефоном в последние три месяца? Пользовались ли Вы смартфоном в последние три месяца?
НН 11	Какой вид (какие виды) интернет-услуг используются для доступа к сети Интернет дома?



НН 12	Как часто Вы обычно пользовались Интернетом в последние три месяца (из любого места)?
НН 13	Пользуются ли данное домохозяйство какой-либо услугой телевидения?
НН 14	Почему у данного домохозяйства нет доступа к сети Интернет?
НН 15	Какие виды деятельности Вы осуществляли в последние три месяца (независимо от вида используемого устройства)?
НН 16	Расходы на ИКТ. Типовой вопрос отсутствует. Данные собираются в ходе обследования бюджетов домохозяйств.
НН 17	С помощью какого устройства Вы пользовались Интернетом за последние три месяца?
НН 18	У Вас есть мобильный телефон? У вас есть смартфон?
НН 19	По каким причинам Вы не пользовались Интернетом?
НН 20	Какие виды товаров или услуг Вы покупали или заказывали через Интернет для личного пользования в последние три месяца?
НН 21	Как Вы оплачивали товары или услуги, которые приобретали через Интернет для личного пользования в последние три месяца?
НН 22	Как Вы получали товары или услуги, которые приобретали через Интернет для личного пользования в последние три месяца?
НН 23	По каким причинам Вы не приобретали товары или услуги через Интернет для личного пользования в последние три месяца?



Несколько десятилетий для измерения социального прогресса, экономического благополучия и производительности рассчитывают и распространяются многочисленные индексы. Индексы, как правило, представляют собой **компаративные показатели**, с помощью которых делается попытка сводной рейтинговой оценки положения отдельной страны на фоне стран мира по разным направлениям.

В их основе лежат данные международных организаций системы ООН (иных объединений стран), которые сформированы при участии национальных статистических систем. В ряде случаев используются базы данных международных организаций, не получившие от Статистической комиссии ООН мандат на формирование и распространение официальной статистики.

Массив исходной информации для таких индексов составляют данные опросов общественного мнения, полученные от выборочных совокупностей, размер которых не превышает 1000 респондентов, в стране а методология формирования выборки описана в общем виде.



**Наиболее известные индексы, характеризующие субъективное благополучие людей приведены ниже:**

1. Better Life Index (OECD) – Better Life Initiative, Index of Happiness, Index of Well-being (индекс лучшей жизни)
2. Legatum Prosperity Index (индекс процветания)
3. Gallup Global Well-Being Index (Индекс глобального благополучия)
4. The Quality of Life Index - (The Economist Intelligence Unit) (Индекс качества жизни)
5. World Happiness Index (мировой индекс счастья).



2023

**WER**  
World Happiness Report

John F. Helliwell, Richard Layard, Jeffrey D. Sachs,  
Jan-Emmanuel De Neve, Lara B. Akinin, and Shun Wang



Для расчета **индекса счастья** наряду со статистическими показателями (ВВП на душу в долларах США по ППС и средняя ожидаемая продолжительность здоровой жизни при рождении) используются ответы респондентов на такие вопросы, как:

- Если вы попали в беду, есть ли у вас родственники или друзья, на которых вы можете рассчитывать на помощь, когда они вам понадобятся, или нет?
- Как ваше физическое и психическое здоровье? - - - Довольны или недовольны вы своей свободой выбирать, что делать со своей жизнью?
- Вы жертвовали деньги на благотворительность в прошлом месяце?
- Распространена ли коррупция в правительстве и в бизнесе?

Метод квантификации субъективных оценок респондента основан на использовании шкалы самоопределения Кантрила.

Оценивается соотношение положительных (смех, удовольствие, интерес) и отрицательных (беспокойство, печаль, гнев) эмоций в течение дня

Методика реконструкции дня).



ПРООН публикует **Индекс общей удовлетворенностью жизнью** измерялся, как среднее значение оценок респондентов по шкале Кантрил. Для расчета индекса используют положительные ответы участников Всемирного опроса Гэллага на вопросы, касающиеся следующих тем:

- удовлетворенность качеством образования;
- удовлетворенность качеством медицинской помощи
- удовлетворенность уровнем жизни;
- ощущение безопасности;
- удовлетворенность свободой выбора;
- восприятие местного рынка труда;
- участие в волонтерской деятельности;
- удовлетворенность общиной;
- доверие к судебной системе;
- удовлетворенность действиями по сохранению окружающей среды;
- доверие к правительству страны.



# **СЕМЕЙСТВО ИНДЕКСОВ РАЗВИТИЯ ЧЕЛОВЕЧЕСКОГО ПОТЕНЦИАЛА, ПУБЛИКУЕМЫХ ПРООН**





### Long and healthy life

Life expectancy at birth



Life expectancy index



### Knowledge

Expected years  
of schooling

Mean years  
of schooling



Education index



### A decent standard of living

GNI per capita (PPP \$)



GNI index



**Human Development Index (HDI)**



### Long and healthy life

Life expectancy at birth



Life expectancy



Inequality-adjusted  
life expectancy index

### Knowledge

Expected years  
of schooling

Mean years  
of schooling



Years of schooling



Inequality-adjusted  
education index



### A decent standard of living

GNI per capita (PPP \$)



Income/consumption

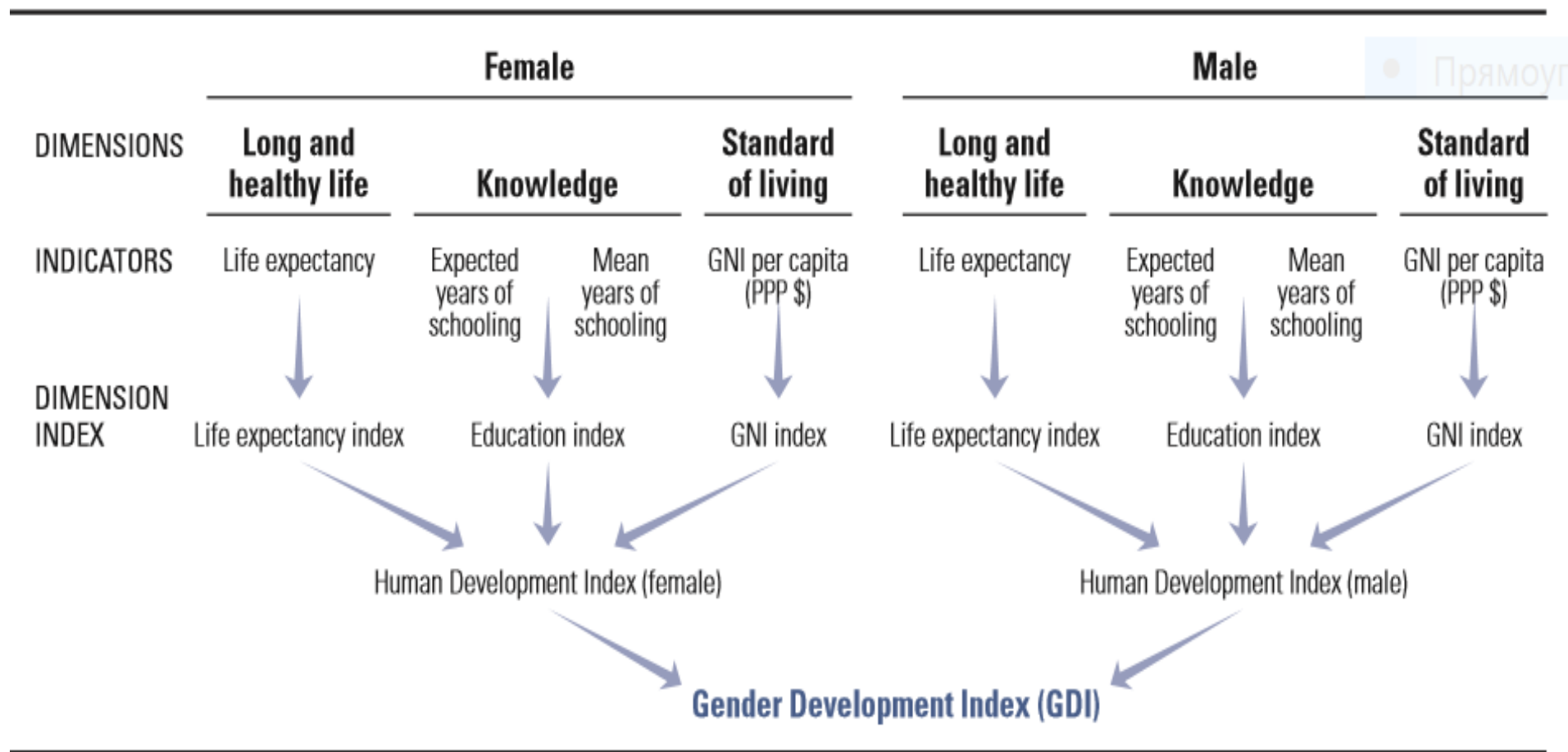


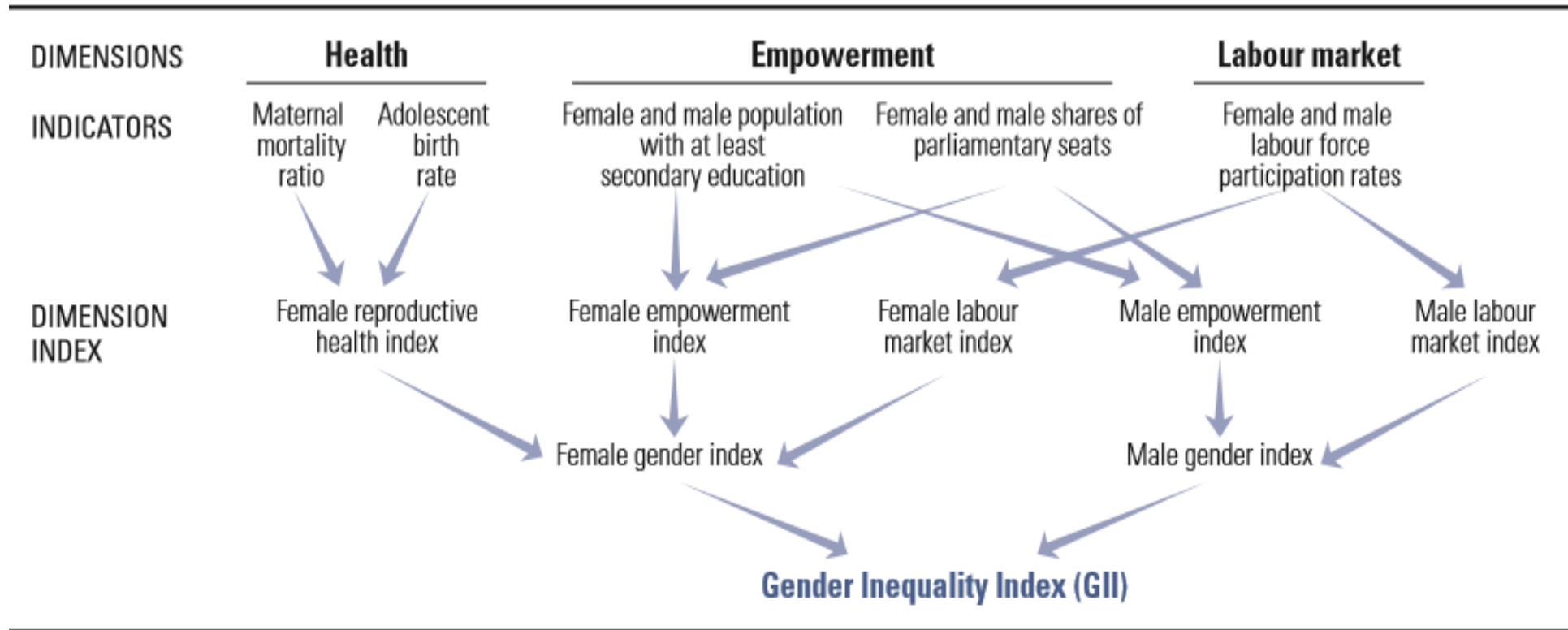
Inequality-adjusted  
income index



**Inequality-adjusted Human Development Index (IHDI)**

Прямоуголь







Для измерения бедности сегодня международные организации применяют многомерные методы, основанные на использовании нескольких деприваций и индикаторов, их характеризующих. Основными игроками в этой области являются ПРООН, Всемирный банк и ЮНИСЕФ (в части детской бедности). Применяемые разными международными организациями критерии и индикаторы для оценки бедности отличаются. Свой подход реализуется в ЕС, где бедность оценивается по показателю риска бедности и социальной исключенности (At Risk of Poverty and Social Exclusion – AROPE).

Области измерения, включенные в расчеты показателей многомерной бедности(число индикаторов)	ПРООН	Всемирный банк	ЮНИСЕФ	ЕС
Здоровье , питание	2		3	1
Образование, учеба	2	2	2	
Условия/ качество жизни	6			
Монетарная мера благосостояния		1		1
Доступность базовых коммунальных услуг		3		
Отдых, досуг, Интернет			7	4
Одежда, обувь			2	2
Низкий уровень занятости				1
Нехватка денежных средств на оплату товаров и услуг, обеспечивающих принятый стандарт жизни				7



## ***Ограничения для сбора статистических данных:***

- ограниченность финансовых ресурсов,
  - нежелание респондентов (и прямой отказ) участвовать в статистических наблюдениях,
  - предоставление респондентами ложных и фрагментарных сведений
  - отсталость технологической основы вычислительной системы НСС
  - ограниченность размеров выборочных совокупностей
  - ограниченность статистических вопросников, не дающая обеспечить получение детализированной информации
  - необходимость обеспечения требований конфиденциальности
  - ведомственный подход к формированию административной статистики и обеспечению доступа к ней
- и др.



## ***Трансформация системы сбора первичных статистических сведений***

предполагает:

- максимально возможный отказ от сбора данных путем организации классических статистических наблюдений;
- применение в статистике сведений, которые отражены в административных записях, то есть прошли проверку при регистрации;  
организация сбора только тех статистических сведений непосредственно от респондентов, которые не могут быть получены из административной статистики и из других источников.

### ***РЕЗУЛЬТАТ***

НСС необременительна для респондентов (и для бюджета):

Респондент не тратит время на участие в статистических наблюдениях, а бюджет – ресурсы на сбор, уже существующих сведений.



## **Перспективы развития источников для официальной статистики:**

- постепенный переход на использование в статистике **административных записей**, регистров, реестров, налоговой, бухгалтерской и финансовой отчетности;
- **ограничение классических статистических наблюдений** сбором данных, которые не фиксируются органами власти;
- расширение программ статистических наблюдений за счет **сбора сведений о фактах со слов респондентов и их мнениях**;
- **использование БД** как источника первичных сведений для официальной статистики





**БЛАГОДАРЮ ЗА ВНИМАНИЕ!**

[surinov@hse.ru](mailto:surinov@hse.ru)