



ОПЫТ РУМЫНИИ в сфере СОПОСТАВЛЕНИЯ ДАННЫХ

Николае Драганеску, Агентство по социальным пособиям и инспектированию, Румыния

Сабин Кирическу, консультант по ИКТ, Всемирный банк

Сессия 6, Румыния

СЕМИНАР

ПО ПРОТИВОДЕЙСТВИЮ ОШИБКАМ, МОШЕННИЧЕСТВУ И
КОРРУПЦИИ (ОМК) В ПРОГРАММАХ СОЦИАЛЬНОЙ ЗАЩИТЫ

8-12 июня 2014 года, г. Опатия, Хорватия



СОДЕРЖАНИЕ

- Зачем
- С чего мы начали
- Выявление несоответствий между данными
- Что мы сделали (2013 г.)
- Усвоенный опыт
- 2014 год, планы на будущее

Зачем использовать сопоставление данных?



Для выявления ситуаций, в которых информация о получателях по-разному внесена в различные государственные базы данных

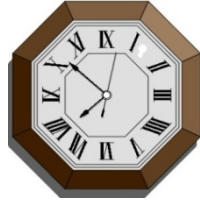
Такое несовпадение данных может использоваться для инспектирования или проведения административных проверок

Такие проверки обходятся дешево по сравнению со сверкой документации на бумаге, имеют комплексный характер и экономически эффективны

- Можно сразу проверить все дела получателей
- Предельные издержки проведения дополнительной проверки близки к нулю
- Высокая эффективность затрат: Вместо выборочных проверок можно целенаправленно проверять дела получателей с несовпадающими данными, особенно по пособиям с высоким уровнем потенциальной переплаты
- Предложения по улучшению совместимости данных между различными организациями для борьбы с ошибками, мошенничеством и коррупцией (ОМК)
- Снижение количества административных ошибок (в информационных системах управления (ИСУ) системы социального обеспечения (СО) и в системах информатизации — внешних источниках данных)

Несовпадение данных → дела, подозреваемые в ОМК → ранжирование по уровню риска → проведение проверки дел с высоким уровнем риска

Контекст — Социальная инспекция в Румынии



- **Социальная инспекция была создана в 2007 году**
- **До 2011 года она занималась социальными услугами**
- **С 2012 года инспекция занимается проверкой программ выплаты пособий путем проведения выборочных проверок, сопоставления данных в ограниченном объеме по программе пособий по инвалидности; при этом следует отметить низкое качество источников данных**
- **2013-2014:** перекрестная сверка данных получателей из ИСУ социального обеспечения (SAFIR) с информацией из других государственных баз данных: службы регистрации актов гражданского состояния, налоговой службы, пенсионной службы, реестров безработных
 - Первое массовое сопоставление данных было проведено в ноябре 2012 — марте 2013 года
 - В процесс было вовлечено 4 масштабных программы социального обеспечения (4,5 млн получателей)
 - Было проверено 11 миллионов записей

Основные программы социального обеспечения, которым проводится сопоставление данных



Программа	Основные критерии участия	Количество получателей
Семейное пособие (СП)	<p>Программа с проверкой нуждаемости: проверяются доходы и имущество</p> <p>Размер пособия: фиксированная сумма на одного ребенка</p> <p>Единица поддержки: семья с детьми</p> <p>Условие: посещение школы</p>	около 260 000
Пособие для малообеспеченных (ПМ)	<p>Программа с проверкой нуждаемости: проверяются доходы и имущество</p> <p>Размер пособия: разница между гарантированным доходом и доходом домохозяйства</p> <p>Единица поддержки: домохозяйство</p> <p>Условия: поиск работы, работа без оплаты</p>	около 170 000
Пособие по уходу за ребенком (ПУР)	<p>Пособие может выплачиваться на протяжении одного года или двух лет.</p> <p>Размер пособия: на первом году 85% среднемесячного дохода, рассчитанного за один год до рождения ребенка; на втором году — фиксированная сумма</p>	около 170 000
Государственное пособие на детей (ГПД)	<p>Пособие, выделяемое на всех детей в возрасте до 18 лет и/или до окончания учебы в школе (лицее или аналогичном учебном заведении)</p>	около 3 800 000

С чего мы начали?

Предпосылки для сопоставления данных



- У каждого гражданина Румынии имеется удостоверение личности с уникальным идентификатором (CNP)
- Большинство государственных баз данных проиндексированы по CNP
- CNP присваивают органы регистрации актов гражданского состояния, которые ведут базу данных населения
- В то же время, базы данных с указанием состава семьи не существует
- В базы данных налоговых органов, органов пенсионного обеспечения, занятости и социального обеспечения вносятся различные виды доходов



С чего мы начали? С идентификации априорных рисков ОМК по каждой из программ



Программа	Риски ОМК	Наличие и качество внешних данных
Семейное пособие (СП) и пособие для малообеспеченных (ПМ) (с проверкой на нуждаемость)	Семья	Отсутствие (достоверных) электронных записей о составе семьи
	Идентификация личности; наличие в реестре населения (например, среди умерших)	Высокое качество данных (служба РАГС)
	Доход (высокий риск)	Высокое качество данных (налоговая и пенсионная службы)
	Имущество (высокий риск)	Отсутствие общенациональных электронных реестров дорогостоящих активов (земельных участков, домов, транспортных средств, скота и т.п.)
	Другие условия: по некоторым из них источники данных существуют, но отсутствуют протоколы сопоставления данных; некоторые из них проверяются через ИСУ (например, наличие прогулов в школе)	
Пособие по уходу за ребенком (ПУР)	Идентификация личности	Высокое качество данных (служба РАГС)
	Доход (высокий риск)	Высокое качество данных (налоговая и пенсионная службы)
	Работа в период получения пособия (высокий риск)	Высокое качество данных (налоговая служба): если отражен доход, то получатель работает, хоть и утверждает, что нет
Государственное пособие на детей (ГПД)	Идентификация личности	Высокое качество данных (служба РАГС)
	Возраст ребенка	Может проверяться; может вызывать ложные подозрения в ОМК

Выявление несоответствий между данными: умерший получатель



Данные о выплатах из ИСУ социального обеспечения (системы SAFIR) (упорядоченные по идентификационному номеру и месяцу)

МЕСЯЦ	ИД. НОМЕР (CNP)	СУММА
Янв 2014	1640403nnnn	100
Фев 2014	1640403nnnn	100
Янв 2014	1750131mmm	100
Фев 2014	1750131mmm	100
Март 2014	1750131mmm	100
Апрель 2014	1750131mmm	120
Май 2014	1750131mmm	120



Информация о рождениях и смертях из реестра

ИД. НОМЕР (CNP)	ИМЯ И ФАМИЛИЯ	ДАТА РОЖДЕНИЯ	ДАТА СМЕРТИ
1750131mmm		31 января 1975 г.	10 марта 2014 г.
123456mmm	Сабин Кирическу	3 апреля 1964 г.	

Выявление несоответствий между данными: Недекларирование какого-либо вида дохода



Для точного сравнения данные о доходах должны быть разбиты по видам, получателям и периодам (месяцам)

МЕСЯЦ	Ид. номер семьи	ИД. НОМЕР (СНР)	ВИД дохода	СУММА дохода
Янв 2014	1	1750131mmmm	Месячная зарплата	200
Янв 2014	1	1750131mmmm	Пенсия	100
Янв 2014	1	2800131nnnn	Аренда жилья	50
Информация о доходах от налоговой администрации				350
Всего по семье				

МЕСЯЦ	ИД. НОМЕР (СНР)	ВИД дохода	СУММА дохода
Янв 2014	1750131mmmm	Месячная зарплата	200
Янв 2014	1750131mmmm	Пенсия	100
Янв 2014	2800131nnnn	Авторские гонорары	50
Всего по семье			350



Выявление несоответствий между данными: Сопоставление суммарных данных по семье!



Иногда бывает сложно сравнивать неидентичные записи или записи с различным уровнем детализации. Есть решение: нужно сравнивать полные доходы семей по данным ИСУ социальной и налоговой служб

МЕСЯЦ	Ид. номер семьи	ВИД дохода	доход СУММА
Янв 2014	1	Месячная зарплата	350
Всего по семье			350

Данные о семье

Ид. номер семьи	ИД. НОМЕР (СНР)
1	1750131mmm
1	2800131nnnn

Информация о доходах от налоговой администрации

МЕСЯЦ	ИД. НОМЕР (СНР)	ВИД дохода	СУММА дохода
Янв 2014	1750131mmm	Месячная зарплата	200
Янв 2014	1750131mmm	Пенсия	100
Янв 2014	2800131nnnn	Авторские гонорары	50
Всего по семье			350

Вывод: эта семья — вне подозрений

Пошаговый подход

Подготовка



1. **Определение риска ОМК (например, недекларирование или неполное декларирование дохода от сдачи имущества в аренду)**

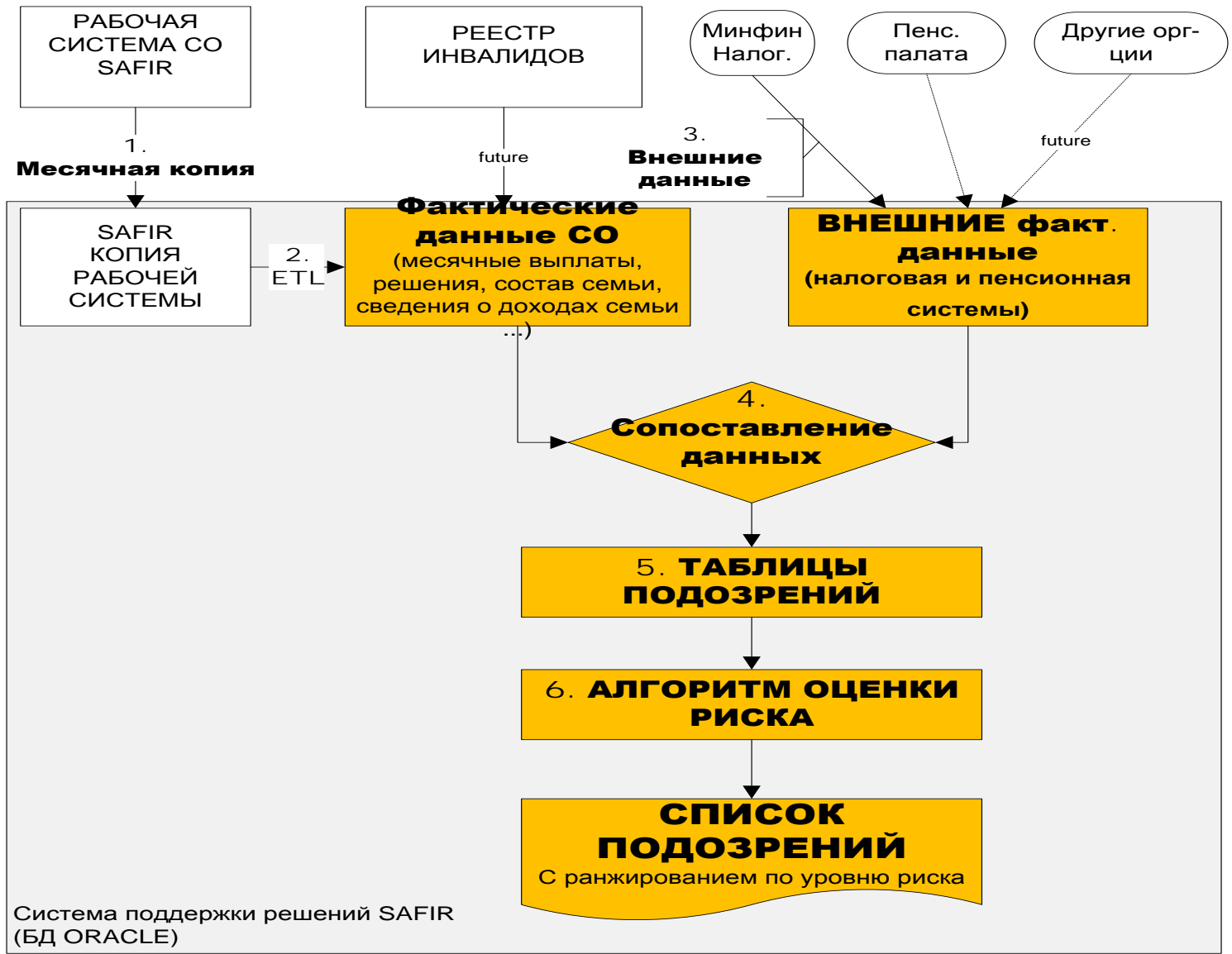
Каких-либо жестких правил выполнения этой задачи нет. Все зависит от требований законодательства, программы социального обеспечения, имеющихся в ИСУ SAFIR данных, а также опыта, накопленного социальной инспекцией

Нужно сосредоточиться на высоких рисках (например, доходах и имуществе)

2. **Есть ли у нас источник внешних данных для сравнения?**
3. **Затем нужно подписать протоколы обмена данными с соответствующим учреждением**
4. **И оценить качество внешних данных (например, данные налоговой службы о доходах, поступающих из-за рубежа, некачественны — они неполны и недостоверны)**

Пошаговый подход

Процесс сопоставления данных



Что мы сделали (2013 г.)

Идентификация рисков — направление проверки



- Проверка **по реестру умерших** для выявления лиц, получающих платежи согласно системе SAFIR, после своей смерти
- Проверка **на предмет несуществующих лиц**, отсутствующих в других базах данных, но получающих выплаты согласно системе SAFIR
- Проверка **доходов** (зарплат и/или пенсий, и/или других социальных выплат), получаемых лицами/семьями одновременно с получением пособий

Что мы сделали (2013 г.)



Данные, которые необходимо получить

- **Личные** данные (идентификационный номер, имя и фамилия, дата рождения, дата смерти)

В Румынии у каждого гражданина имеется личный идентификационный номер (CNP), используемый во всех основных общегосударственных базах данных

- **Данные о доходах** (зарплате, пенсии, пособиях, арендной плате, роялти, авторских гонорарах и т.п.)

Учреждения, предоставляющие данные

- Министерство государственных финансов, Национальная налоговая администрация,
- Национальное агентство занятости,
- Служба реестра населения — МВД
- Национальная государственная пенсионная палата
- **Сама ИСУ SAFIR для получения информации о доходах в виде различных пособий (на детей, на уход за ребенком — суммы этих пособий добавляются к доходу семьи при назначении СП и ПМ)**

Что мы сделали (2013 г.)

Источники данных



ВИД ДАННЫХ	ПРЕДОСТАВЛЯЮЩЕ Е УЧРЕЖДЕНИЕ
Идентификационные данные	Налоговый центр (а не реестр населения)
Месячный доход (зарплата, пособие по безработице — всего около 20 видов)	Налоговый центр
<i>Годовой доход (арендная плата, роялти, авторские гонорары)</i>	<i>Налоговый центр</i>
Месячный пенсионный доход	Пенсионная палата
Месячный доход от социальных пособий	NASBSI - сама система SAFIR

Вывод: все дело в протоколах и формате обмена данными

Что мы сделали (2013 г.)

Подозрение — что мы ищем



(программы: СП, ПМ, ПУР, ГПД)

➤ Подозрения в отношении «личных данных»:

- Несуществующие лица
- Умершие
- По возрасту (дети) — только в отношении ГПД

➤ Подозрения в отношении «доходов»:

- **Завышение доходов в случае ПУР** для увеличения суммы пособия (т.к. на протяжении первого года сумма пособия составляет 85% задекларированного среднемесячного дохода)
- **Недекларирование (или неполное декларирование) доходов в случае ПМ**

Что мы сделали (2013 г.)



Использование данных с учетом потребностей и ограничений в сфере социального обеспечения

- Мы сравнивали неидентичные данные, несмотря на риск ложных сигналов путем сравнения **оценки** чистого дохода (по данным налоговой службы о заработной плате до уплаты налогов и правилам налогообложения) с **фактической** чистой зарплатой, указанной в ИСУ социального обеспечения

Для расчета чистого дохода использовался алгоритм, предусмотренный законодательством о социальном обеспечении

- Мы фильтровали данные путем перекрестной проверки дел с наиболее высокой вероятностью нарушений

Урок:

- во избежание слишком большого количества «ложных срабатываний» в обеих базах данных должны быть ИДЕНТИЧНЫЕ данные
- Тем не менее, неидентичные данные все равно позволяют выявлять дела с высоким риском ОМК.

Лучше начать сопоставление данных уже сейчас, чем ждать идеальные данные

Что мы сделали (2013 г.)

Ранжирование подозрений по уровню тяжести



- Сформированный список подозрительных дел явно выходил за пределы возможностей социальной инспекции (из-за сопоставления оценочных показателей с фактическими)
- В связи с этим была проведена оценка уровня риска по каждому из подозрительных дел
 - по шкале от 1 до 10 (наиболее высокий риск)
 - Уровень риска был выше в случае большой продолжительности нарушений, большой суммы переплаты или значительных расхождений (между задекларированной зарплатой) между обеими базами данных
 - Социальная инспекция сосредоточилась на рассмотрении дел с наиболее высоким уровнем риска
 - После этого по результатам проведенных проверок определялось, сработал ли наш алгоритм оценки риска

Примеры параметров для пособия ПУР (риск)

- Длительность переплаты в месяцах
- Общая сумма пособий за период переплаты
- Размер среднемесячного дохода до уплаты налога согласно внешним данным

Повышенный уровень риска (9 или 10) в случае большой продолжительности переплаты (более 6 месяцев), значительных сумм переплаты и высоких доходов до уплаты налога

Определение уровня риска играет важную роль в определении приоритетности проверки различных дел

Что мы сделали (2013 г.)

Пример перекрестной проверки (СП)



ИД. НОМЕР СЕМЬИ	ВИД ПОДОЗРЕНИЯ	Длительность ПЕРЕПЛАТЫ В МЕСЯЦАХ	УРОВЕНЬ РИСКА	ИД. НОМЕР (СНР)	ИМЯ И ФАМИЛИЯ	РАБОТОДАТЕЛЬ ID	ПОСОБИЕ	Доход по данным НАЛОГОВОЙ службы	ПЕНСИОННЫЙ доход	Доход от СОЦИАЛЬНЫХ ПОСОБИЙ
nnn	Лимит СП	6	7	1	Родитель 1	10953691		600	-	-
						13341930		191	-	-
						16892658		2,471	-	-
						14075542		594	-	-
				2	Ребенок		ГПД	-	-	630
				3	Родитель 1			-	5 625	-
СОВОКУПНЫЙ ДОХОД за ПЕРИОД ПЕРЕПЛАТЫ								3 856	5 625	630
СОВОКУПНЫЙ ДОХОД СЕМЬИ								10 111		
СРЕДНИЙ ДОХОД СЕМЬИ ЗА МЕСЯЦ								1 685		
СРЕДНИЙ ДОХОД НА ЧЛЕНА СЕМЬИ ЗА МЕСЯЦ								562		
ЛИМИТ ДОХОДА для СЕМЕЙНОГО ПОСОБИЯ по результатам проверки на нуждаемость (на члена семьи)								370		

Что мы сделали (2013 г.)



Сводные результаты сопоставления данных, 2013 год

Программа	Направление работы	Количество проверенных дел	Количество проверенных лиц	Количество платежей, подлежащих проверке	Количество дел с подозрением на высокий уровень ОМК	Количество дел с подозрением на низкий уровень ОМК	Общее количество дел, подозреваемых на наличие ОМК
ПУР 26 месяцев: Янв 11 — Фев 13	Получатель	435 178	785 801	4 713 983	544	5 719	6 263
	Доход	435 178	785 801	4 713 983	23 612	6 943	30 555
ПМ 15 месяцев: Дек 11 - Фев 13	Получатель	282 874	544 096	2 978 448	3 048	3 137	6 185
	Доход	143 739	168 043	1 379 948	2 080	7 699	9 779
СП 15 месяцев: Дек 11 - Фев 13	Получатель	425 232	1 370 455	4 503 198	7 426	4 899	12 325
	Доход	295 074	933 863	3 074 041	14 405	24 934	39 339
ГПД 26 месяцев: Янв 11 — Фев 13	Получатель	4 463 455	7 890 364	100 057 521	9 544	2 404	11 948
	Возраст	4 463 455	4 411 719	100 057 521	50 970	478 584	529 554 <i>Причиной большого количества дел с потенциальным и проявлениями ОМК может быть неполнота данных</i>

Что мы сделали (2013 г.)

Общие результаты сопоставления данных



	Количество подозрительных дел — сопоставление данных	Количество дел, изученных в 2013 году	Количество дел с ОМК, 2013 год	Общая сумма переплаты (задолженность) 2013	Уровень взыскания задолженности к 31 марта 2014 года	Затраты на 31 декабря 2013 года
ПУР	36 818	27 720	5 650	11 569 170	5 117 824	115 692
СП	51 664	34 123	15 151	2 643 421	1 690 999	1 129 939
ПМ	15 964	9 676	2 959	1 730 223	877 075	
ГПД	541 502	27 308	2 549	1 189 902	692 476	74 964
Отопление	35 851	24 030	4 874	755 368	241 046	490 989
Инвалидность	5 457	5 457	2 724	1 301 069	551 914	143 118
Итого		128 314	33 907	19 189 153	9 171 334	1 954 701

Соотношение «затраты/эффект» по переплатам (оценка объема задолженности)

9,58

Соотношение «затраты/эффект» по взысканной задолженности

4,69

Уроки, вынесенные из первой попытки сопоставления данных



- До 2013 года не существовало ни концепции, ни возможности перекрестного сопоставления данных за длительный период времени, хранимых в системе SAFIR и системах других государственных органов
- В центре внимания в ходе контроля были подозрения на ОМК → с целью экономии значительных материальных, финансовых и человеческих ресурсов
- Изменение подхода к контролю с → чисто административного на проведение проверок при наличии подозрений

Работа по сопоставлению данных оказалась первой работой такого масштаба (десятки миллионов строк, три источника данных)

Уроки, вынесенные из первой попытки сопоставления данных



- Данные различных органов между собой не гармонизированы (до или после уплаты налогов; разные типы доходов)

В 2014 году в налоговое законодательство по просьбе социальной и статистической служб были внесены изменения, предусматривающие указание чистых доходов.

- Сопоставление данных и формирование профиля рисков начинается с изучения критериев программы (если не запрашивать доход на человека и в разрезе видов дохода с самого начала, то проверить доходы позже будет гораздо труднее)
- Данные нужно понимать. ➔ Если понимания нет — будет много подозрительных дел
- Очень важно определять уровень риска ➔ СИ должна рассматривать наиболее подозрительные дела в приоритетном порядке

Сопоставление данных 2013



Незначительные ресурсы — Отличная команда —
Высокий результат

**Половина команды по сопоставлению
данных (треть группы информатизации)**

Они занимаются не только сопоставлением данных

**Рабочее место для
сопоставления данных
Без ноутбуков**



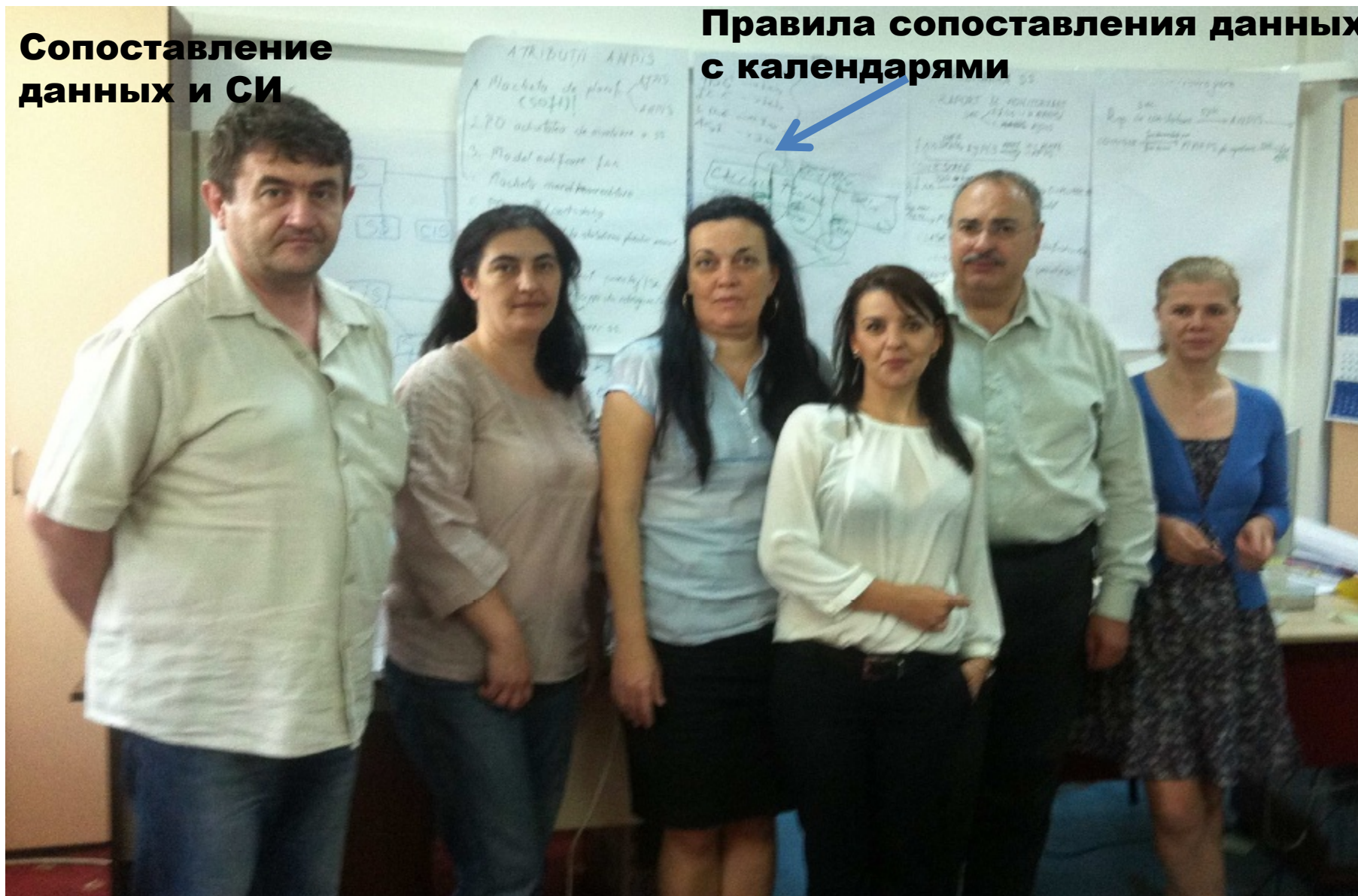
Сопоставление данных 2013

Незначительные ресурсы — Отличная команда —
Высокий результат



**Сопоставление
данных и СИ**

**Правила сопоставления данных
с календарями**



2014 год: сопоставление данных становится стандартной практикой работы



- 2012 год: первые попытки сопоставления данных (не в массивах, без стандартизации)
- 2013 год: первое массовое сопоставление данных по 6 программам
- **2014 год: Сопоставление данных проводится дважды в год; оно стало обычной функцией Национального агентства социальных выплат**
- Пути совершенствования:
 - Новые критерии перекрестной проверки: Средний доход по ПУР для расчета суммы пособий
 - Новый подход: проверка семей, а не отдельных людей → в этом случае дело можно проверить один раз на предмет всех возможных нарушений (до 2014 года СИ работала по программам)
 - Усовершенствованный алгоритм расчета уровня риска (критерии и их вес)

Планы расширения сопоставления данных



- Крупные программы (СП, ПМ, ПУР, ГПД) проверяются дважды в год
- Пособия по инвалидности и на отопление проверяются раз в год
- Выявление новых источников данных для дополнительных перекрестных проверок
 - Годовой доход (арендная плата, роялти, авторские гонорары)
 - Безработица — ежеквартальное подтверждение права на получение пособия (“adeverinta ANOFM”)
- Выявление новых групп риска и видов нарушений
- Разработка специализированного инспекционного модуля ИСУ

Специализированный инспекционный модуль ИСУ — Сквозная система мониторинга



- Планирование проверок
 - Сбор и регистрация сведений о подозрительных делах (ОМК) по разным каналам (общественность, персонал, сопоставление данных, профилирование риска)
 - Документирование формирования выборки дел для проверки
- Проведение проверок
 - Регистрация результатов каждой проверки
 - Регистрация принятых мер, примененных санкций, а также сумм переплаты или недоплаты
- Отчетность о результатах и непрерывный мониторинг
 - Формирование таблицы с основными результатами проверки в целях мониторинга
 - Контрольная проверка через несколько месяцев

Планы на будущее

Перспективный процесс взаимодействия ИСУ и социальной инспекции

