

А. В. Кузнецов

Племучет: ошибки в идентификационных записях импортных животных

Аннотация. Приведена классификация искажений информации, встречающейся в племенных записях. Дано описание генезиса искажений. Приведены примеры информационных коллизий из официальных источников: баз данных и экспортных сертификатов. Показана необходимость создания национальной базы сведений о племенных животных федерального значения. Обоснована ее арбитражная функция. В статье разъяснен ряд новых терминов. Табл. 2 и рис. 5.

Ключевые слова: племенной учет, ошибки в племенных записях, официальный источник данных, национальная база сведений о племенных животных, арбитраж.

Автор:

Кузнецов Андрей Васильевич — кандидат сельскохозяйственных наук, доцент, консультант ОАО «Краснодарское» по искусственному осеменению с.-х. животных, директор Союза «Некоммерческое партнерство животноводов Краснодарского края», Краснодар, тел. 8 (918) 112-44-77, e-mail: 89181124477@mail.ru.

Селекционеры всегда официально заявляют, что ошибки в племенных записях недопустимы. И тихо добавляют: «Но они там есть...»

Целью является разработка системы выявления и исключения ошибок в племенных записях. В задачи исследования входит теоретическое обоснование и классификация ошибок, встречающихся в племенных идентификационных записях, а также демонстрация примеров типов ошибок из официальных зарубежных источников.

Содержание понятия «идентификационная запись» до сих пор обсуждается специалистами. От функционального наполнения этого понятия зависит содержание официальных документов, удостоверяющих племенной статус животных. Нами в рамках исследования принято, что идентификационная запись о животном — это совокупность сведений, которые:

- по своей природе являются качественными;
- известны, установимы либо распространяют свое действие с момента рождения животного;

- не требуют специальных методик и инструментов для определения параметра;
- не могут измениться в течение всей жизни животного.

Согласно этим требованиям к идентифицирующим признакам относятся: кличка животного, индивидуальный номер (международный идентификационный номер), дата рождения животного, порода и породность животного, его племенной номер (номер регистрации в национальной племенной системе), а также клички и индивидуальные номера родителей животного.

Любое изменение идентификационной записи следует считать ее искажением (рис. 1). Искажения информации делятся на два типа: ошибки записи и коллизии, каждый из которых делится на два вида.

В рамках исследований слово «коллизия» (от лат. *collisio* — столкновение) означает в равной степени подтвержденные юридически, но отличающиеся друг от друга идентификационные записи

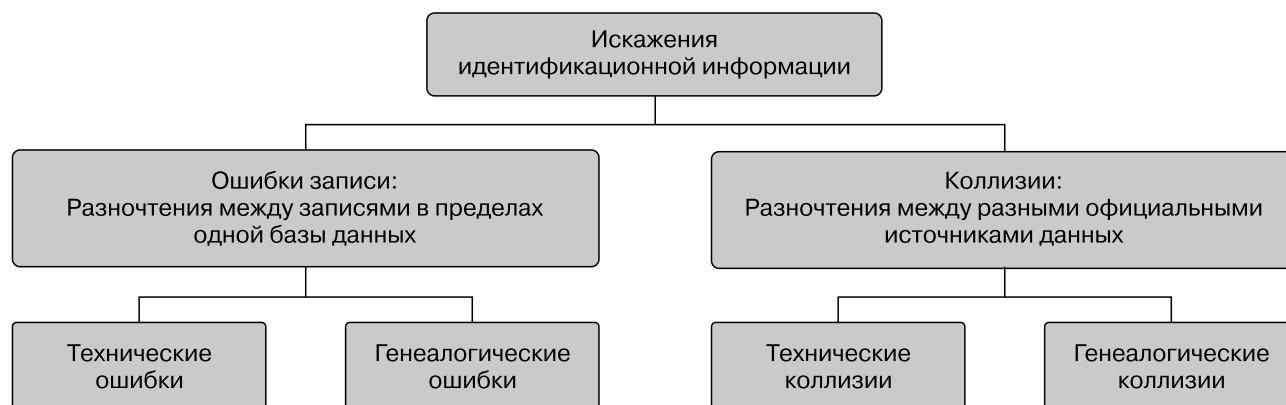


Рис. 1. Классификация искажений идентификационных записей о животных

одного и того же животного. Отсюда принципиальное значение имеет анализ процесса возникновения, срока «жизни» ошибок, а также момент трансформации ошибок в коллизии (Рис. 2).

Чем отличаются друг от друга типы искажений: ошибки и коллизии?

1. Ошибки создаются, тогда как коллизии обнаруживаются. Следовательно, нельзя понять — как долго существует коллизия, и как широко она распространена среди информационных реципиентов проверяемого официального источника.

2. Ошибки существуют в пределах одной базы данных, коллизии — в нескольких официальных источниках. Следовательно, исправить ошибку можно, а коллизию — практически невозможно,

так как каждый модератор официального ресурса, скорее всего, будет отстаивать свою версию, как им проверенную.

3. Как показывает практика — создателя ошибки установить возможно. Следовательно, можно понять причины модификации информации и оценить необходимость этой модификации. Отсюда зависит направленность поиска ошибок того же типа. В то же время создателя коллизии установить практически невозможно. Следовательно, невозможно установить мотивы и обоснованность существования той или иной коллизии.

Отсюда следует, что если практика исправления ошибок может осуществляться в пределах одной производственной базы данных, то практика

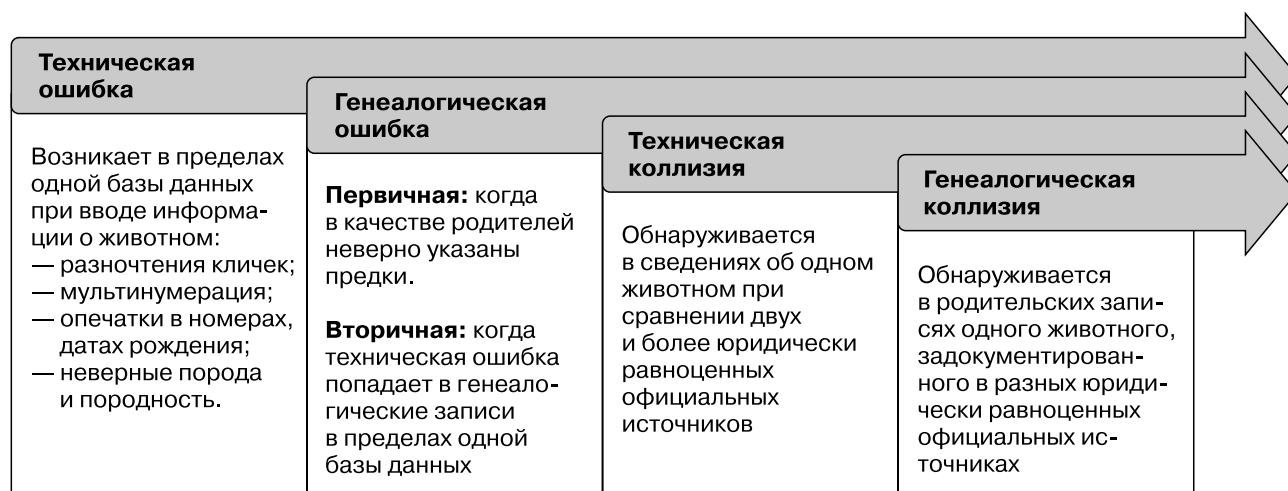


Рис. 2. Генезис искажений идентификационных записей о животных

Таблица 1. Неоднозначная идентификационная информация о быке
ETAZON LABELLE NLD 460942030

Ресурс	Страна	Кличка	Ассоциированные номера
www.cdn.ca	Канада	ETAZON LABELLE	HONLDM460942030 0097HO00002
www.holsteinusa.com	США	528 ETAZON LABELLE-ET	USA 2247421
www.holstein.ca	Канада	ETAZON LABELLE PB	HONLDM460942030
www.accelgen.com	США	528 Etazon Labelle-ET	HOUSA000002247421 097HO000002
www.crv4all.com	США	ETAZON LABELLE	NLDM000460942030
www.holstein.com.au	Австралия	ETAZON LABELLE-MFF-BLC-CVC	NLDM460942030 A00009455, NLDLABELLE
www.fabaweb.mloy.fi	Финляндия	Etazon Labelle Bl	F 92456 C
www.adhis.com.au	Австралия	ETAZON LABELLE	A00009455, NLDLABELLE
www.icbf.com	Ирландия	ETAZON LABELLE ET	ELB HOLNLDM000460942030
www.wwsires.com	США	LABELLE*BL	97HO00002 HOUSA000002247421
www.vikinggenetics.fi	Сев. Европа	E Labelle	NLD000000460942030 FIN 92456, SWE 99719, DNK 234277
www.dairybulls.com	США	528 ETAZON LABELLE-ET	USAM000002247421 097HO00002

одновременного исправления коллизий в ряде официальных баз данных, вероятней всего, обречена на неудачу.

Для оценки объемов работы по приведению коллизий идентификационной информации к одному знаменателю нами собрана информация о достаточно часто встречающемся быке-производителе по 12 официальным базам данных (табл. 1).

Модификации клички быка являются типичным примером технической коллизии. В данном случае человеку понятно, что речь идет об одном производителе. Но это не понятно компьютеру. В итоге модификация клички может привести к клонам записей в одной базе данных (рис. 3)

Бесспорно, что ИАС СЕЛЭКС справляется с учетными операциями на фоне существования большого количества клонов записей. Так что с точки зрения информационной безопасности проблема в системе ИАС СЕЛЭКС решена. Но есть проблема зоотехнической целостности информации. И она постоянно заявляет о себе при подборе производителей к стадам, при оформлении племенных документов, обмене репликами баз данных между хозяйствами и регионами и т.д. Принято по умолчанию, что ее решение — не функция ИАС СЕЛЭКС, а обязанность селекционера. Но для комплексного решения данной проблемы необходимо снабдить селекционера соответствующим инструментом, имеющим достаточный юридический вес для доказательства его правоты. Такого инструмента в РФ пока нет.

Существует мнение, что идентификационные потребности полностью исчерпываются индивидуальным номером животного. Однако не следу-

ет думать, что кличка является третьестепенным идентификационным признаком. Достаточно сказать, что в отличие от индивидуального номера кличка несет еще и некую семантическую составляющую, уменьшающую утомляемость специалиста при оперировании большим количеством данных. Кроме того кличка легче поддается проверкам. Известны ситуации, когда не индивидуальный номер, а именно кличка является уникальным идентифицирующим признаком (Рис. 4).

Сложнее обрабатывать коллизии индивидуального номера. В табл. 1 имеется 9 позиций, где указаны индивидуальные номера и идентификаторы, присвоенные быку Etazon Label в разных национальных племенных системах. Каждый из номеров — равноценный официальный идентификационный признак. Поэтому, если в Россию попадут экспортные сертификаты из США (www.holsteinusa.com) и Австралии (www.holstein.com.au) — то, скорее всего, селекционер задокументирует в одном случае американский, а в другом случае австралийский номер, а компьютер расценит эти варианты, как описывающие разных животных.

На самом деле в разных базах данных разные номера одного животного нередко связаны между собой. Ввод любого из них вызывает появление одной конкретной записи. Такую связь предлагаем называть ассоциацией, а номера — ассоциированными. Наличие ассоциативной связи между номерами позволяет:

- проверить правильность записи обоих номеров и связанных с ними данных;
- корректно сопоставить родословные и потомство;

Инв.№	Кличка	Дата рожд.	Код пор.	Порода
122358313	Мэн		32	Ч/п голштинская
122358313	О-Мен	08.03.1998	32	Ч/п голштинская
122358313	О Мен Джустис	08.03.1998	32	Ч/п голштинская
122358313	Манфред Джустик		32	Ч/п голштинская
122358313	Манфред Джустик		32	Ч/п голштинская
122358313	джустис	08.03.1998	32	Ч/п голштинская
122358313	Мэн.	08.03.1998	32	Ч/п голштинская

Рис. 3. Результаты запроса быка в ИАС СЕЛЭКС по индивидуальному номеру 122358313

Name	Phys ID	Sex	Herd Book#	Birth Date	Class	NASIS
ELMHILL STATESMAN GAY		Female	55582	18/07/1964	G78 14/06/1970	
KENROSS KING MACK		Male	55582	05/01/1976		
NOTRELLE SOVEREIGN IRENE		Female	55582	16/06/1981		
TAHORA REVENUE LEYTON		Male	55582	13/05/1975		LEYTON

Рис. 4. Результат запроса по номеру 55582 на сайте www.holstein.com.au (Австралия)

— установить ассоциации третьих номеров.

Однако сводные таблицы ассоциированных номеров — большая редкость. Наиболее распространены пары «номер — NAAB», реже «номер — NASIS». На сайте www.vikinggenetics.fi официально представлены ассоциированные группы номеров «международный — финский — шведский — датский». Последний пример — единственный пример таблицы ассоциированных номеров, публикуемый в свободном доступе. Тем не менее, мультиноммерация часто встречается у зарубежных животных (табл. 2).

В российской системе племенного учета нередки номера-ошибки, возникшие в результате опечаток, либо каких-то соглашений селекционеров. Например, выявлен зарегистрированный в базе ИАС СЕЛЭКС номер 1650414502383 быка S-W-D VALIANT. На самом деле этот номер образован слиянием американского и немецкого номеров. Мотивы нам неизвестны.

Часто встречаются намеренные сокращения индивидуальных номеров у быков. Зоотехнические службы объясняют это неудобством работы с длинными номерами. Но в итоге «сокращенный» номер не пригоден для поиска животных в сторонних информационных системах, уменьшает степень индивидуализации животного при поиске.

Гораздо реже технических коллизий встречаются генеалогические коллизии. Их наличие ставит под сомнение не только идентификационную запись потомка и/или родителя, но и всю линию предков спорного родителя.

Например, из пакета австралийских сертификатов нами выделены 4 документа, в которых быку FATAL FRA 2290038601 сопоставлено 4 разных матери (рис. 5). Проверка происхождения быка в австралийской базе данных показала, что на самом деле ему сопоставлена мать BRUNE 84 FRA 2286037388. Это же подтвердил аналогичный запрос на сайте www.cdn.ca.

Таким образом, фактически зоотехник обязан внести в базу данных животное, которое не значится в экспортных сертификатах. Но сегодня нет нормативного документа, позволяющего игнорировать коллизии в официальных источниках данных, к каковым относятся и экспортные сертификаты.

Неочевидные коллизии — один из сложнейших случаев в зоотехнической практике верификации учетных данных. Установление действительно верной записи иногда не может опираться на сведения лишь из одной базы данных. Для устранения коллизий нужен арбитраж по нескольким официальным базам данных. Но и он не всегда дает однозначные сведения.

На наш взгляд, выходом из ситуации — и искомым инструментом доказательства правоты селекционера — является создание национальной базы сведений о племенных животных, работающих на территории Российской Федерации. Ее назначение:

- верификация идентификационных сведений из сторонних официальных источников;
- аккумуляция идентификационных сведений племенных животных Российской Федерации;

Таблица 2. Примеры мультиноммерации быков-производителей в мировой практике племенного учета

Кличка	Номер 1	Номер 2	Номер 3	Номер 4	Номер 5	Номер 6
TO-MAR BLACKSTAR-ET	USA 1929410	DEU 502870	SWE 99185	DNK 226201	FIN 91717	DEU 10502870
ROUND OAK RAG APPLE ELEVATION	USA 1491007	DEU 502043	SWE 99587	DNK 82084	DEU 10502043	AUS A00005151
ETAZON CELSIUS	USA 2247437	DEU 501319	SWE 90022	DNK 233463	FIN 92700	NLD 460508522
S-W-D VALIANT	USA 1650414	DEU 502383	SWE 99091	DNK 82097	FIN 90070	AUS A00005152
CARLIN-M IVANHOE BELL	USA 1667366	DEU 502528	SWE 90016	DNK 17001	FIN 91670	AUS A00005425
ETAZON ADDISON-ET	NLD 839380546	DEU 501551	SWE 90239	DNK 241230	FIN 92448	AUS A00011809
SKALSUMER SUNNY BOY	NLD 311651443	DEU 501257	SWE 99709	DNK 228603	FIN 92287	AUS A00008144
BIS-MAY TRADI- TION CLEITUS	USA 1879085	DEU 502797	SWE 99146	DNK 223803	FIN 91850	NLD 970020790
CAL-CLARK BOARD CHAIRMAN	USA 1723741	DEU 502578	SWE 99105	DNK 17002	FIN 91673	AUS A00004775

Выдержки из экспортных сертификатов Австралии (указаны номер бирки, номер чипа телки; номер сертификата; запись о предке с вариациями)

Certificate serial number : AH2007110123944 Visual ID : G0015 Electronic Tag : 982 000038491643 Date of Birth : 16/08/2005	Name : FATAL HB No : FRA2290038601 Australian Breeding Values : 08/2007 Milk : 158 Fat : -12 Prot : 18 Rel : 99	Name : UGELA BELL HB No : FRA4483049978 Australian Breeding Values : 08/2007 Milk : 508 Fat : 9 Prot : 21 Rel : 74 Name : TARINKA PARK MASCOT DEANNE HB No : AUS627364
Certificate serial number : AH2007110112859 Visual ID : P0484 Electronic Tag : 982 000060743734 Date of Birth : 12/08/2006	Name : FATAL HB No : FRA2290038601 Australian Breeding Values : 08/2007 Milk : 158 Fat : -12 Prot : 18 Rel : 99	Name : UGELA BELL HB No : FRA4483049978 Australian Breeding Values : 08/2007 Milk : 508 Fat : 9 Prot : 21 Rel : 74 Name : GLOMAR MANHATTAN LUCY HB No : AUS721365
Certificate serial number : AH2007110106091 Visual ID : P0550 Electronic Tag : 982 000060237608 Date of Birth : 20/07/2006	Name : FATAL HB No : FRA2290038601 Australian Breeding Values : 08/2007 Milk : 158 Fat : -12 Prot : 18 Rel : 99	Name : UGELA BELL HB No : FRA4483049978 Australian Breeding Values : 08/2007 Milk : 508 Fat : 9 Prot : 21 Rel : 74 Name : COMESTAR LOR BLACK HB No : CAN5311227
Certificate serial number : AH2007110144436 Visual ID : Q5208 Electronic Tag : 982 000086550902 Date of Birth : 03/06/2006	Name : FATAL HB No : FRA2290038601 Australian Breeding Values : 08/2007 Milk : 158 Fat : -12 Prot : 18 Rel : 99	Name : UGELA BELL HB No : FRA4483049978 Australian Breeding Values : 08/2007 Milk : 508 Fat : 9 Prot : 21 Rel : 74 Name : ELITE SOUTHWIND BELLADONNA HB No : AUS507380

Сведения из официальной базы данных Австралии

[www.http://abri.une.edu.au/online/cgi-bin/i4.dll?1=223D2B2A&2=2420&3=56&5=2B3C2B3C3A&6=5D5D24595B582321&9=5054](http://abri.une.edu.au/online/cgi-bin/i4.dll?1=223D2B2A&2=2420&3=56&5=2B3C2B3C3A&6=5D5D24595B582321&9=5054)

— Sire: UGELA BELL (PENN-SPRINGS MR C)-BLC-CVC FRAM 4483049978 VG

Animal: FATAL-BLC-CVC (g) FRAM 2290038601

— Dam: BRUNE 84 FRAF2286037388

Рис. 5. Ошибка в указании матери быка, выявленная на этапе сличения родословных экспортных сертификатов из Австралии

- отслеживание движения племенных животных;
- свободное предоставление информации любому заинтересованному лицу.

Весь процесс сличения проблемных родословных с референтной базой достоверных племенных регистраций, ранее подтвержденных перекрестным анализом, предлагается называть аутентификацией родственных записей (аутентификация от греч. αὐθεντικός [authentikos] — реальный, подлин-

ный) в зоотехническом значении слова «сличение с подлинной родственной записью с целью исключения различий».

Использование национальной базы сведений о племенных животных должно носить арбитражный характер при верификации генеалогических записей племенных животных. Аккумуляция в данной базе не только генеалогических записей, но и ассоциированных с ними регистрационных данных (в том числе ошибочных номеров) поз-

волит создать арбитражный инструмент, главной целью которого будет исключение неоднозначных трактовок происхождения и идентификации животных или так называемых клонов. Это позволит мотивированно исключать спорные случаи

при анализе информации о племенных животных, опираясь лишь на один отечественный источник, а не на множество зарубежных, и одновременно упростит весь процесс доказательства правоты селекционера.



Kuznetsov A.

Pedigree records: errors in identification records of imported animals

Abstract. *The article provides the classification of misrepresentations of data in pedigree records. The genesis of misrepresentations has been described. Examples of information collisions have been given according to the data from official sources: databases and export certificates. The paper shows the necessity of developing the National pedigree database of federal significance. The arbitration function of the database has been substantiated. The article explains a range of new terms. Tabl. 2 and Fig. 5.*

Key words: pedigree records, errors in pedigree records, official data source, National pedigree database, arbitration.

Author:

Kuznetsov A. — PhD (Agr. Sci), associate professor, Adviser of the JSC «Krasnodarskoe» for artificial insemination of farm animals, Director of the Union «Non-profit partnership of livestock breeders of Krasnodar region», Krasnodar, tel. 8 (918) 112-44-77, e-mail: 89181124477@mail.ru.