An Integrated Database

Mike Lynch
Database Analyst, ICBF
ICAR Cork 2012

Session T3: Information Systems
ICBF – an integrated database

- Overview
- Data Sources
- Data Set
- Services Provided
- Technologies
ICBF Database

- Entire Irish cattle breeding data set
- Highly integrated database
- Data from many sources
- Multiple organisations provide data:
  - Herd books
  - Milk Recording organisations
  - AI companies
  - Farming Organisations
  - Dept. of Agriculture
ICBF Database

- Organisations involved
  - Provide data
  - Use database services
    - Data Input
    - Reports
    - Pedigree Certificate
    - Farmer Services
    - Online Services

- Large number of individual end users
  - Farmers
  - Advisors
  - Vets
  - Marts
Data Sources

SERVICES

Genomics

Health

Factories

Linear Scoring

AI Recording

Milk Recording

Animal Events & Farmer Recording

Animal Movements

Calf Birth Data
Core Data Sets

Animal

Sex  Birth Date  Sire  Dam
Location  Breed
Milk  Health  Weight  Fertility
Measurement  Genetic Evaluations
Size of Regular Data Set

- 2.6 million milk records per year
- 7 million movements per year
- 2 million calf births per year
- 81,000 herds
Size of Larger Data Sets

- 2.3 Billion Genotype SNP Records
- 1.1 Billion Breeding Value Records
Database Growth

Total Animal Records (millions)

- Suckler Cow Welfare Scheme
- AI Companies
- Beef Herd Books
- Holstein Friesian & Milk Recording

© Irish Cattle Breeding Federation Soc. Ltd 2012
Accessing the Database

- 2001 - Server – Client Model
  - Third Party Client Software
  - Organisations accessed WAN over ISDN
- 2003 - Intranet Services
  - Increased services using intranet developed in-house
- 2005 - Introduction of Internet Services
  - Farmers & Advisors could access reports online
- 2006 - Herdplus Service
  - Enhanced commercial web service to farmers
- 2008 - Herd Book Web App. (developed in-house)
  - Herd books move away from legacy system
- 2009 - Introduction of MRO Web Application
  - MROs move away from legacy system
- 2010 - More organisation types with web access
  - Enhanced services for vets, advisors, consultants etc.
Web Servers

Display

Web Server

Operating System

Virtualisation
ICBF Network
October 2011

- Two sites connected by fibre.
- Internet connection to each provides failover.
- Cisco switches and ASA.
- HP-UX servers.
- Dell Servers connected to HP EVA 6100 storage running VMWare.
ICBF Development Team

- **Team**
  - 7 full time developers
  - 2 System Administrators
  - 2 Handheld Computer Specialists
  - 1-2 contractors

- **Responsibilities**
  - Database Software Development
  - Web Development
  - Document Production
  - Handheld Software Development
  - Routine Data Transfers
  - Database Administration
  - System Administration
Integrated Database - Key Features

• Clean Data Set
  – Avoid Duplication
  – Data should be as clean as possible at source

• Animal Identification is vital
  – Single Animal ID works best

• Use reliable transfer methods
  – Look at speed, reliability, security
  – Do no over complicate
  – Try to use same methods where possible

• Access
  – Database is only useful if people can access it
  – Internet Technology has become faster, more reliable and is highly adaptable
Future Areas of Interest

• Very Large Data Sets
  – How to develop sustainable tools/systems to deal with the volumes of genotype data expected?
  – Data Mining, “Big Data”

• International collaborative projects
  – Optimir, IgenoP

• Handheld Development
  – Data Recording Hand Helds – future growth
  – Smartphone Apps – iPhone, Android

• Virtualisation
  – To what extent do we move to virtual environments
  – What platform do we build on (VM, Microsoft, etc)
  – Cloud Computing