Employing high resolution big data for predictive modelling in precision dairy farming

G. Katz

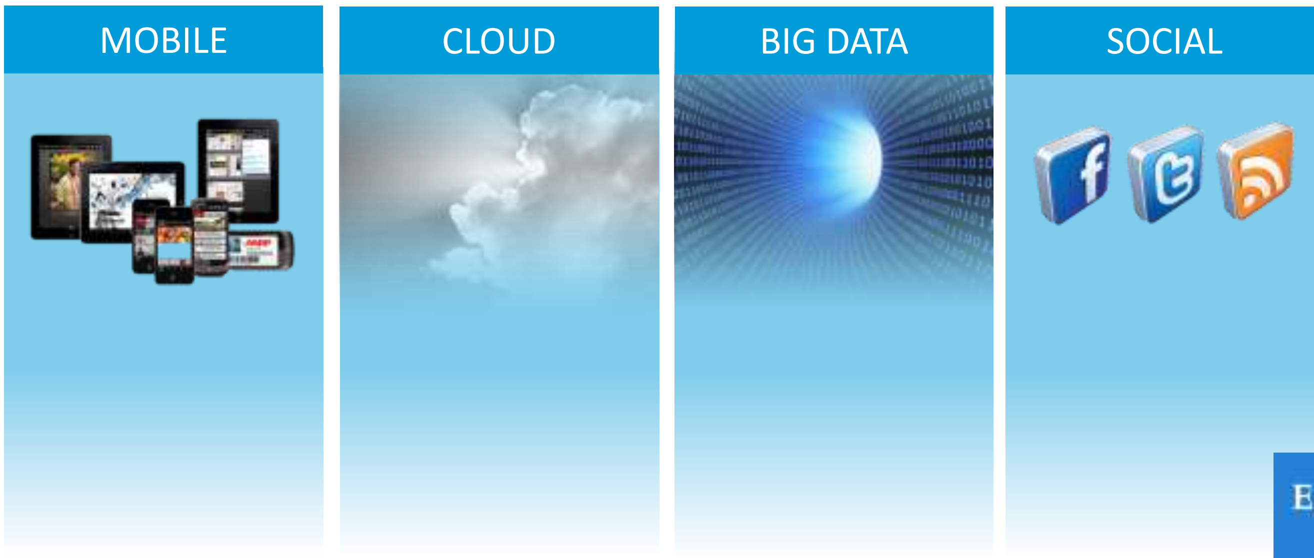
Speaker: Gil Katz

Dairy farming in the emerging era of IOT

Gil Katz

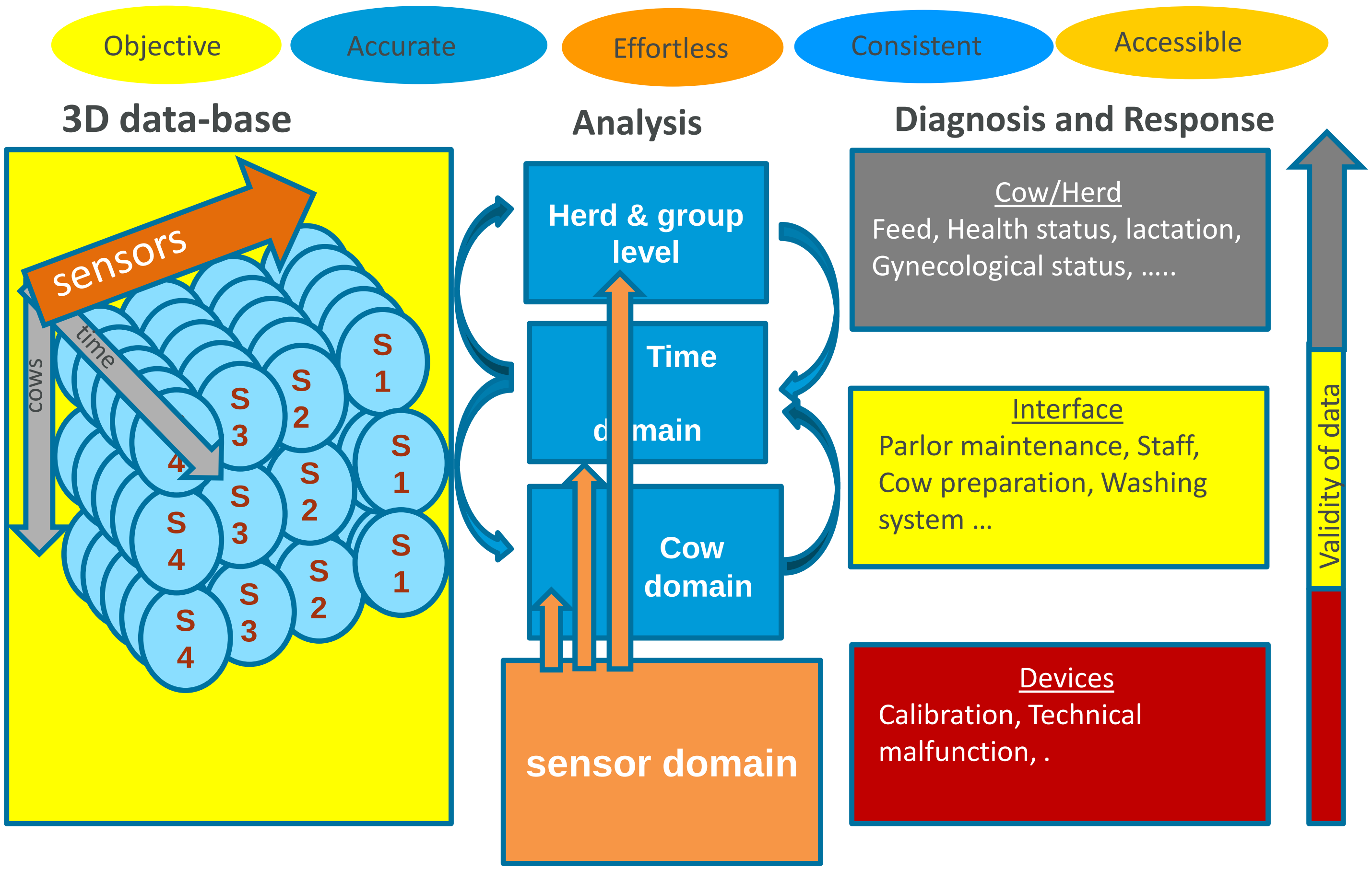
Afimilk

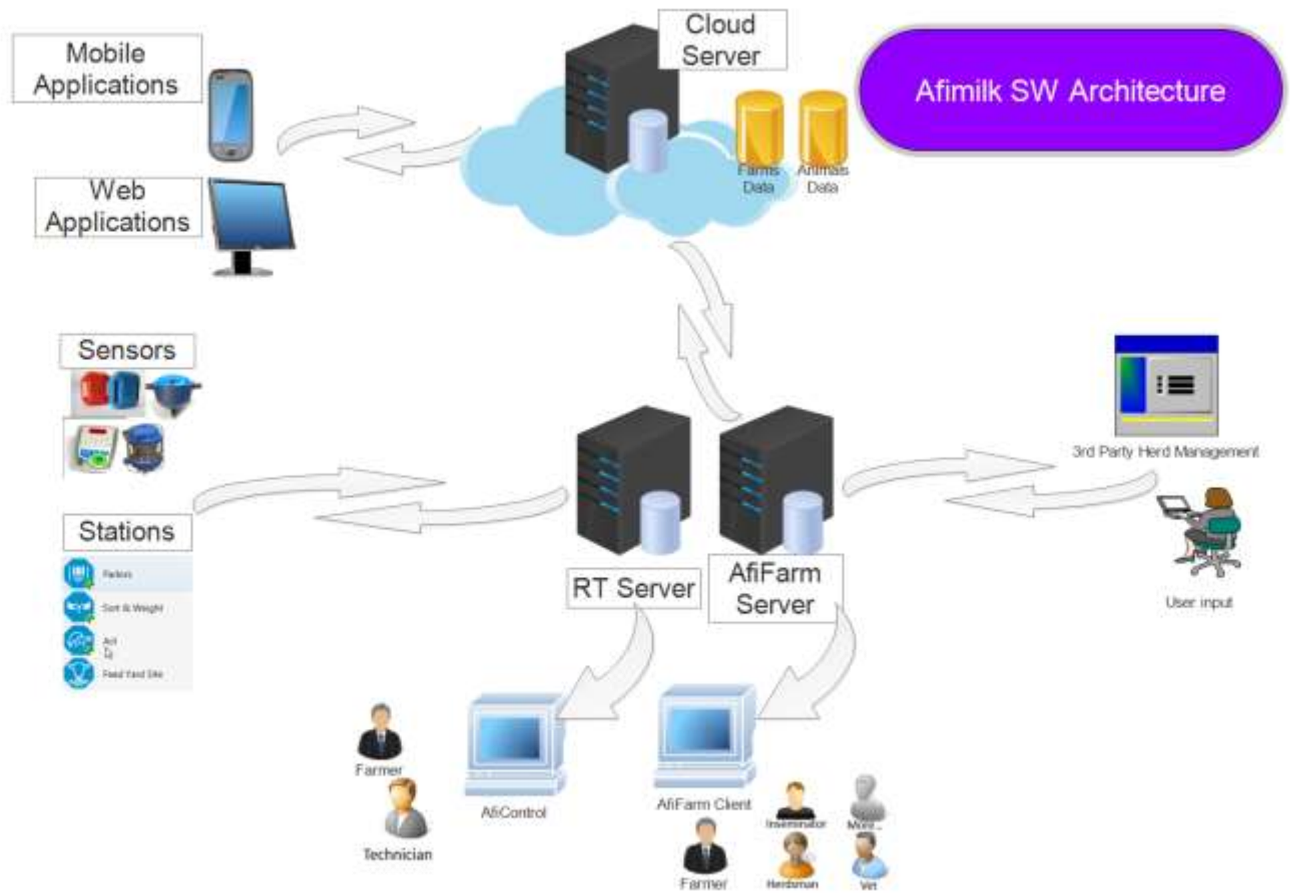
CONVERGENCE OF MEGA TRENDS



The INTERNET OF THINGS

Automated Data Collection and Analysis




Afimilk SW Architecture
Mobile Applications
Web Applications
Cloud Server
Farms Data
Animals Data
Sensors
Stations
Sort & Weight
Feed Yard Site
RT Server
AfiFarm Server
3rd Party Herd Management
User input
Farmer
Technician
AfiControl
AfiFarm Client
Inseminator
More...
Farmer
Herdsman
Vet

Manage and merge different Data types

Quantitative data (monotonic structure)
milk yield, milk components, milk flow, weight ….

Qualitative data (discreet structure)
gynecological status, health status …

Behavioral data (pattern based)
activity pattern, grouping pattern, rest pattern, feed pattern ….

Milking stall sensors – milk yield, milk flow, milk conductivity, milk fat, protein, lactose, blood, coagulation potential

Cow sensors – activity, lying times, lying bouts

Big Data

Complex biological systems

Challenge:
construct data, collect data, mine data, Develop predictive models, Validate models, construct comparative standards

Disciplines

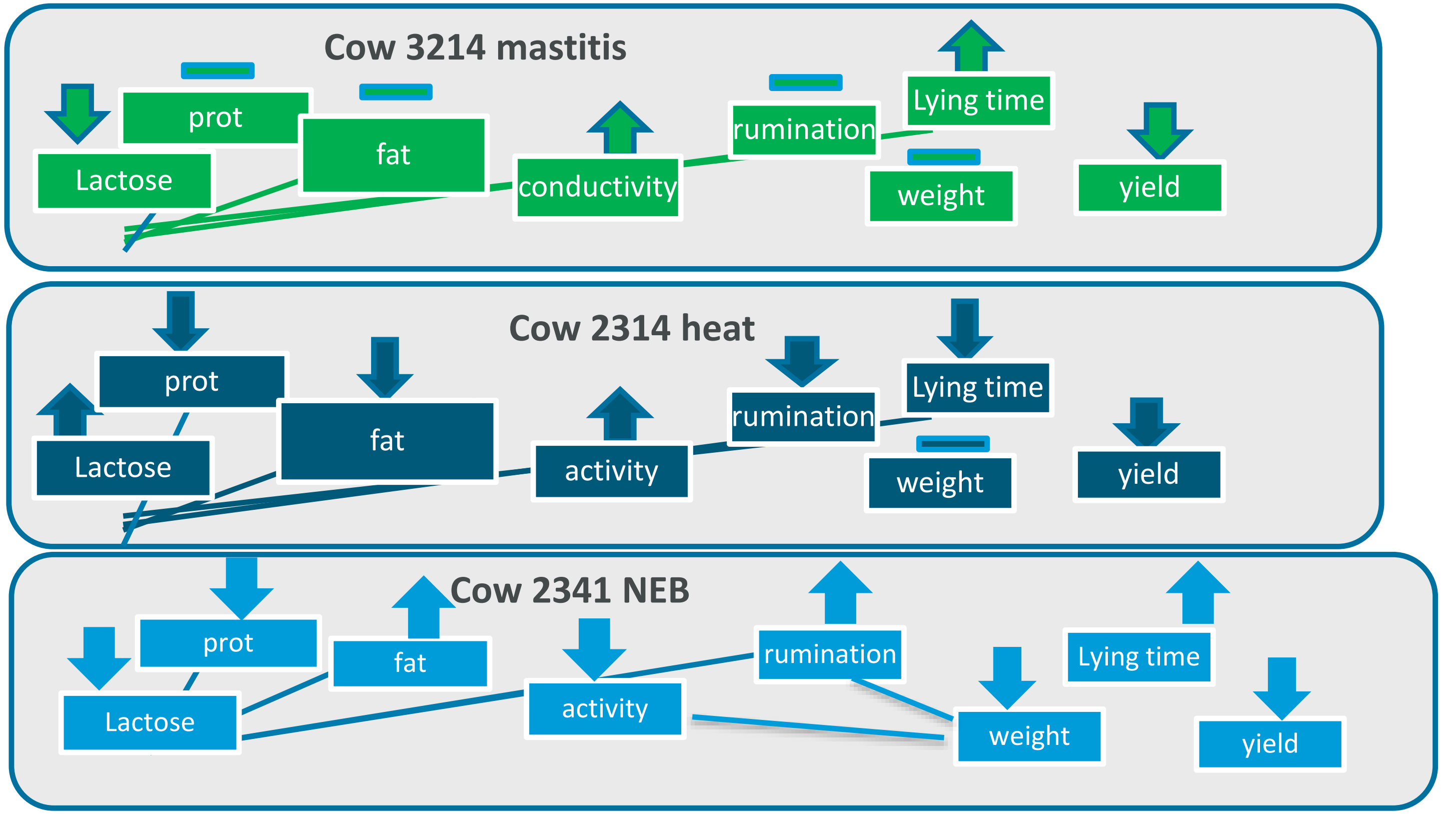
Biology, Chemistry, Physics

Data science, Mathematics, computer science

Health, Fertility, Feed, Genetics, Production

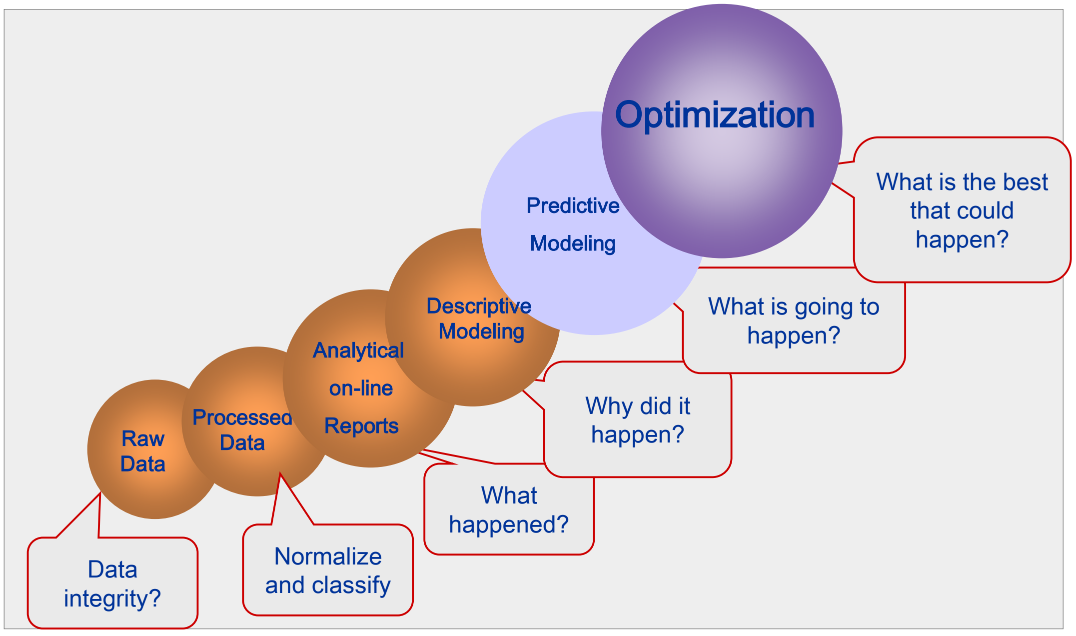
Challenge: Pattern recognition of subjective multi dimensional data

Descriptive :From highlighting irregularities to diagnostics



From Data Collection to Decision Making

Data → Information → Knowledge → Intelligence



Arkadi Slezberg, 2009

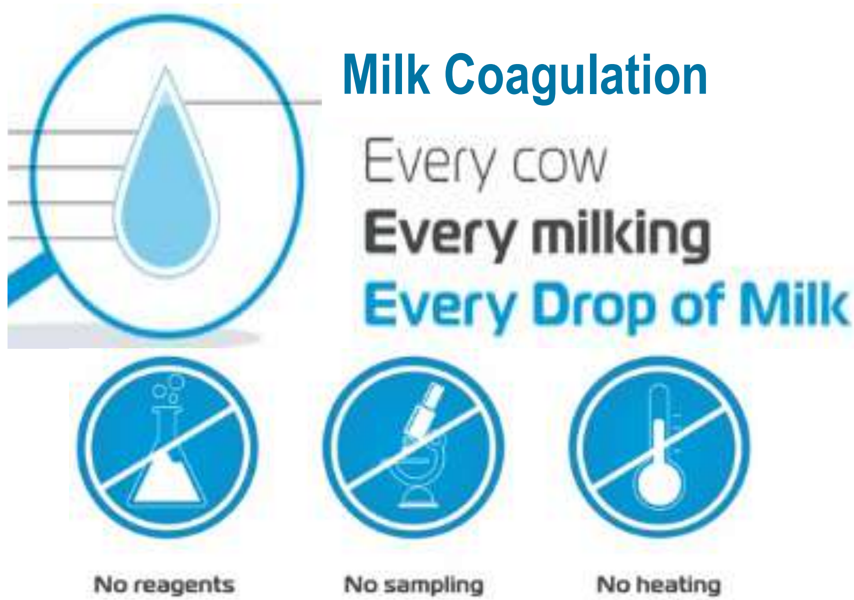
From retrospective to prospective prediction of production

Real time measurement of milk yield and composition

AfiLab concept

Blood
Lactose
Protein
Fat

Milk Coagulation



No additives

No cost per sample

No farther procedures

- **Casein, un-saturated fatty acids, saturated fatty acids, mono & poli Unsaturated fatty acids , igG count in colostrum**

Predictive : From diagnotics to prediction

Different heuristic approach

From classical statistics terminology:

- Mixed models
- Decision trees
- Bayesian models

To time dependent terminology:

- Dynamic modeling
- Markovian and non-Markovian processes
- Memory stamps

J. I. Weller and E. Ezra, "*Genetic and phenotypic analysis of daily Israeli Holstein milk, fat, and protein production as determined by a real-time milk analyzer*", JDC, Vol. 99 No. 12, 2016

- Scope: >37,000 Holstein cows spanning over 2 years
- Finds agreement between Afimilk's inline milk lab real time analysis and between DHIA monthly tests.
- Selected for 'Editor's Choice‘ of JDSc

Objectives of the study

- ✓ Comparison of lactation yields between the traditional testing & Afilab
- ✓ Calculation & comparison of Predicted Transmitting Ability (PTA)
- ✓ Calculation of genetic & phenotypic correlations
- ✓ Establishing correction factors for Season, Age & Open Days
- ✓ Calculation of extended yield factors for cows with truncated data (partial records)

Heritabilities, genetic and environmental correlations among 7,866 first parity 305 d lactations computed from the ICBA and AfiLab records.

Trait	Heritabilities		Correlations	
	ICBA	AfiLab	genetic	environmental
Milk (kg)	0.33	0.35	1.00	0.96
Fat (kg)	0.23	0.31	0.59	0.70
Protein (kg)	0.27	0.32	0.86	0.87
% fat	0.48	0.57	0.70	0.66
% protein	0.55	0.46	0.56	0.52

Heritabilities were higher for the AfiLab records for all traits, except for % protein.

Phenotypic correlations among complete and extended 1st parity lactations the last ICBA test day and the last two weeks of AfiLab records.

FAT (kg)									
Trait	Mean days in milk at truncation								
	30	60	90	120	150	180	210	240	270
ICBA	0.67	0.75	0.79	0.87	0.91	0.93	0.95	0.95	0.96
Afilab	0.77	0.84	0.89	0.92	0.94	0.95	0.96	0.96	0.97

PROTEIN (kg)									
Trait	Mean days in milk at truncation								
	30	60	90	120	150	180	210	240	270
ICBA	0.70	0.76	0.78	0.87	0.90	0.92	0.94	0.94	0.95
Afilab	0.72	0.83	0.87	0.90	0.93	0.94	0.95	0.95	0.96

SUMMARY Weller & Ezra

- ✓ The genetic values for 1st lactation cows were higher by Afilab except for % protein
- ✓ The prediction coefficients for 305 days Kgs milk, fat & protein were higher for Afilab
- ✓ The genetic & phenotypic correlations to 305 days lactation in 30 DIM are 0.75 and gradually rising to 0.98 in 240 DIM
- ✓ Prediction of complete lactation yields from partial data were more effective in Afilab

Prediction of complete lactations in Afifarm

- Our objective: To adapt the **large scale retrospective** study's method to a **prospective prediction** of complete (305_days) lactations in individual herds
  - ✓ For selection
  - ✓ For production planning (quota, summer/winter)
- The operational need: To enable farmers to get the decision as early as possible, **but before breeding**

Oded Nir (Markusfeld)

Waiting Periods

Herds	Cows/herd	Voluntary waiting period (days)	Days to 1st AI
13,885	158.4 ± 325 SD	58.4 ± 5.6 SD	95.2 ± 26.9 SD

Ferguson J.D. & Skidmore A. (2013). JDS 96 (2) 1269 -1289

Days to 1st AI	50	51 - 80	81 - 110	111 - 150
1st lactation	0.4%	41.4%	45.2%	13.0%
2nd lactation	9.7%	58.4%	26.9%	5.1%

Ezra E. (2013). HerdBook Summary (Hebrew). ICBA

Our objective is to be able to make the decision at 60 DIM

Oded Nir (Markusfeld)

Predictive : From diagnotics to prediction

Early prediction of total lactation performance

Prediction calculated from 2014 data (new) compared to 2015 data (old)

- Calibration of models from cows calving in 2014 (26/01-31/12)
- Validation of models applied cows calving in 2015
- 6 herds of Israeli Holsteins with 371 to 1046 annual calving events and 11,840 Kg to 13,635 annual milk

Criteria for Success

- R^2= RSquare of the **summary of fit**
- r = **Correlations** to actual production
- 75% & 90%tiles of **the differences between the predicted & actual estimates** of the various traits **(for planning & selection)**
- **Predictive Values & accuracy for selection decisions**
  - ✓ PPR (positive predicting value)=The probability that a cow defined by test as a "low yielder" is truly so
  - ✓ NPR (negative predicting value)=The probability that a cow defined by test as a "high yielder" is truly so

Oded Nir (Markusfeld)

Afimilk; Herd #3

	Milk, kg/305 days			Fat, kg/305days			Protein, Kg.305 days			ECM, kg 305 days		
	34	54	84	34	54	84	34	54	84	34	54	84
RSquare	0.683	0.726	0.786	0.704	0.737	0.704	0.653	0.698	0.768	0.717	0.753	0.804
Correlations	0.930	0.949	0.968	0.926	0.931	0.926	0.918	0.935	0.956	0.923	0.941	0.962
+tive PV	65.0%	72.2%	84.6%	47.5%	57.6%	47.5%	65.0%	80.0%	84.6%	52.9%	56.7%	76.5%
-tive PV	78.6%	79.3%	79.0%	86.1%	88.4%	86.1%	78.6%	78.7%	79.0%	83.3%	82.6%	81.0%
Accuracy	75.0%	77.6%	80.0%	65.8%	75.0%	65.8%	75.0%	78.9%	80.0%	69.7%	72.4%	80.0%
10%tile to 90%tile	-10.1% to 8.4%	-7.5% to 9.2%	-4.7% to 8.6%	-11.4% to 7.0%	-9.5% to 6.8%	-11.4% to 7.0%	-8.7% to 9.8%	-7.1% to 10.1%	-4.0% to 9.0%	-11.8% to 4.6%	-9.3% to 6.3%	-5.5% to 7.0%

Herd #3: n for 12/14-11/15=717 (34 DIM); 1,195 (54 DIM); 1,912 (84 DIM); n for 12/14-02/16=76

- **Prediction of all the production variables examined improved with time from calving**
- **The smaller herd behaved similar to the larger one**

Oded Nir (Markusfeld)

Afilab <=34 DIM vs. 1st ICBA milk test <=34 DIM (All lactations combined)

	Milk, kg/305 d		Fat, kg/305 d		Protein, Kg.305 d		ECM, kg 305 d	
Herd #1	Afi	ICBA	Afi	ICBA	Afi	ICBA	Afi	ICBA
RSquare	0.568	0.554	0.523	0.388	0.543	0.502	0.571	0.513
Correlations	0.858	0.800	0.866	0.727	0.845	0.784	0.860	0.777
+ve PV	75.0%	54.2%	60.6%	40.9%	71.4%	66.7%	75.0%	57.1%
-ve PV	83.1%	79.1%	87.0%	71.1%	82.8%	76.9%	83.1%	78.3%
Accuracy	81.0%	70.1%	75.9%	61.2%	79.7%	74.6%	81.0%	71.6%
10%tile to 90%tile	-9.3% to 10.3%	-10.4% to 10.7%	-10.8% to 6.8%	-14.3% to 9.8%	-9.9% to 8.7%	-12.2% to 11.2%	-9.4% to 9.9%	-9.7% to 12.3%

Prediction for milk & fat, proved superior to that of ICBA (truncation at 34 DIM)

Oded Nir (Markusfeld)

Afimilk; Afilab + Predicted Transmitting Ability (PTA}
All lactations combined. Herd #3

	Milk, kg/305 days		Fat, kg/305days		Protein, Kg.305 days	
	DIM34	+PTA	DIM34	+PTA	DIM34	+PTA
RSquare	0.683	0.782	0.704	0.744	0.653	0.719
Correlations	0.930	0.942	0.926	0.927	0.918	0.935
+tive PV	65.0%	75.0%	47.5%	51.4%	65.0%	63.6%
-tive PV	78.6%	86.5%	86.1%	87.2%	78.6%	79.6%
Accuracy	75.0%	82.9%	65.8%	69.7%	75.0%	75.0%
10%tile to 90%tile	-10.1% to 8.4%	-10.2% to 5.4%	-11.4% to 7.0%	-11.1% to 9.7%	-8.7% to 9.8%	-8.1% to 7.1%

Adding PTA to the 34 DIM models in Herd #3 proved contributed more than in the 54 DIM models. Results were not different in Herd # 1

Oded Nir (Markusfeld)

Summary & Conclusions

- ✓ Prospective prediction of complete lactations in individual herds yielded similar results to Weller & Ezra's large retrospective study
- ✓ Predictions using Afimilk in 34 DIM proved superior to those using the first Milk Test
- ✓ Though prediction improves with time in lactation, the present results allow for "safe" selection, culling & production planning at 54 DIM, and even earlier in lactation.
- ✓ *Results for small & large sized herds were similar*
- ✓ *Current average production planning error based on ICBA data is 20%-25% using daily afilab data the error drops down 5%-7%*
- ✓ Adding PTA to the models slightly improved prediction of milk & protein in early lactation

Oded Nir (Markusfeld)

Take home message:

Not using available Daily data is a drawback to the industry.

Data reduction by averaging it is loss of information and knowledge.

Thank you