Andreas Holzinger

VO 706.996 W DiplomandInnen Seminar
VO 706.997 W DissertantInnen Seminar

Welcome to the Holzinger Group HCI-KDD
Part 1: What is the HCI-KDD approach and what is aML vs. iML?

a.holzinger@hci-kdd.org
http://hci-kdd.org/scientific-working-for-students

A word at the very beginning

- The “best” is the enemy of the “good” – whenever you try to be “perfect” – there is the danger that you finalize nothing*) …”

*) zero, nada, null

François-Marie Arouet (1694 – 1778)
known as “Voltaire”

Motto of the Holzinger Group

<u>Science</u> is to test crazy ideas – <u>Engineering</u> is put these ideas into Business!

Learning Goals

- At the end of the whole seminar you should
- … be aware of the HCI-KDD approach
- … know the interests of the Holzinger group
- … have an overview on possible research topics
- … be familiar with the formal PhD requirements
- … understand how to carry out scientific research
- … know how to write scientific papers
- … most of all: getting started with your work

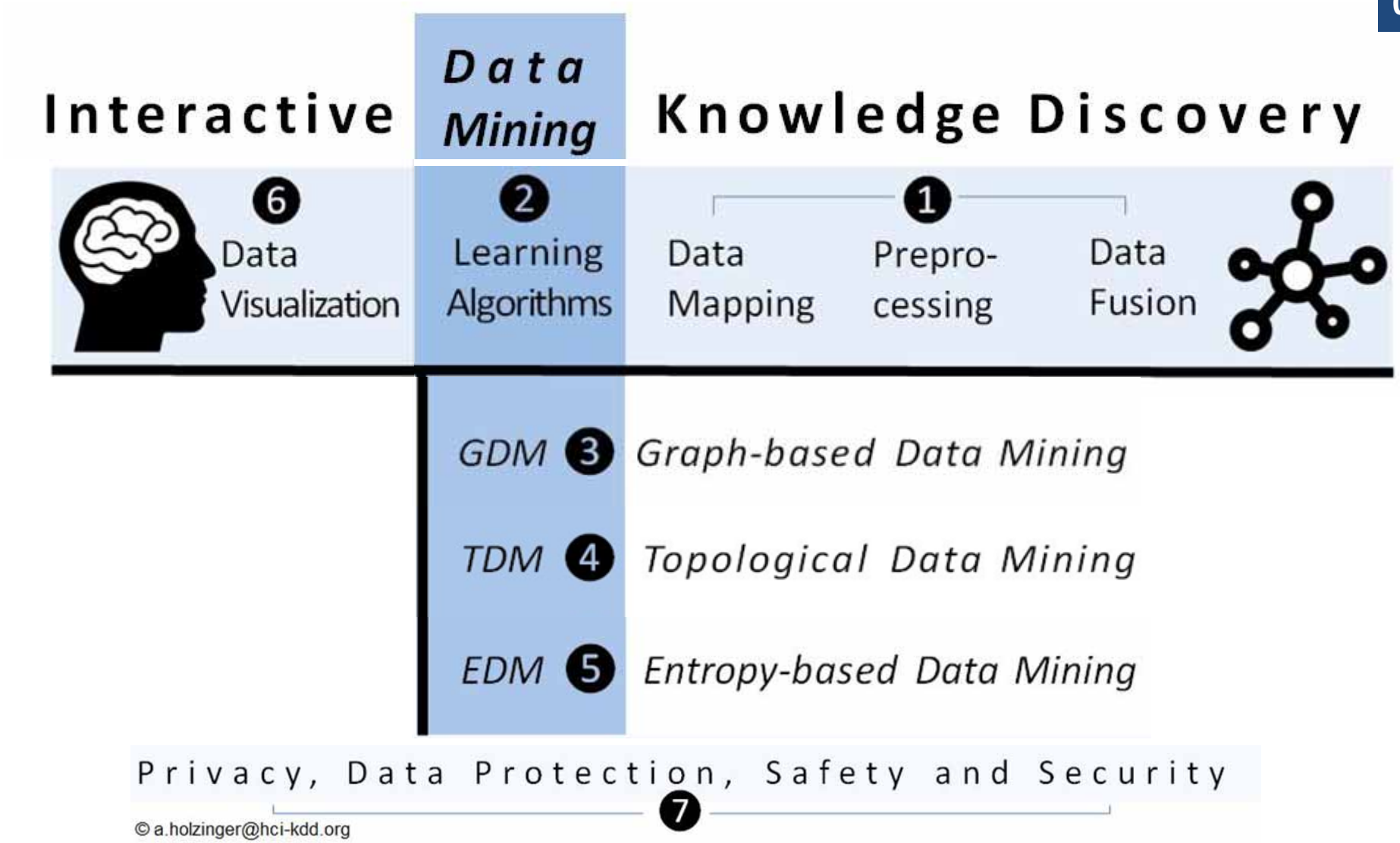
Agenda for lecture 01

- 01 What is the HCI-KDD approach?
- 02 Application Area: Health
- 03 Probabilistic Information
- 04 Gaussian Processes
- 05 Automatic Machine Learning (aML)
- 06 Interactive Machine Learning (iML)
- 07 Conclusion

01 What is the HCI-KDD approach?

- Machine Learning is a very practical field – at the core is **algorithm development** – however, successful machine learning needs a concerted effort of various fields …

The heart in our Knowledge Discovery pipeline is ML …

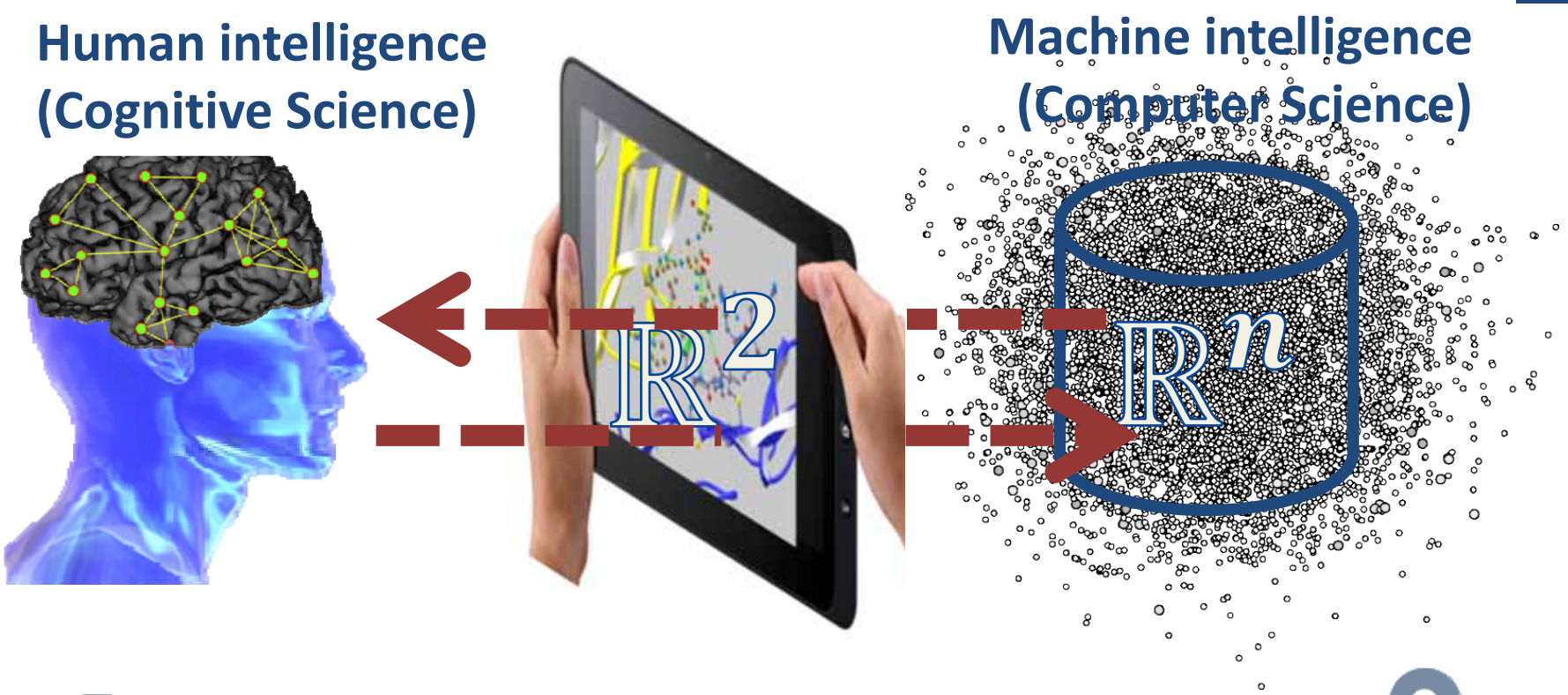


Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning. IEEE Intelligent Informatics Bulletin, 15, (1), 6-14.

… but successful ML needs a concerted effort!


http://hci-kdd.org/international-expert-network

Grand challenge: Transfer results from $\mathbb{R}^n$ to $\mathbb{R}^2$



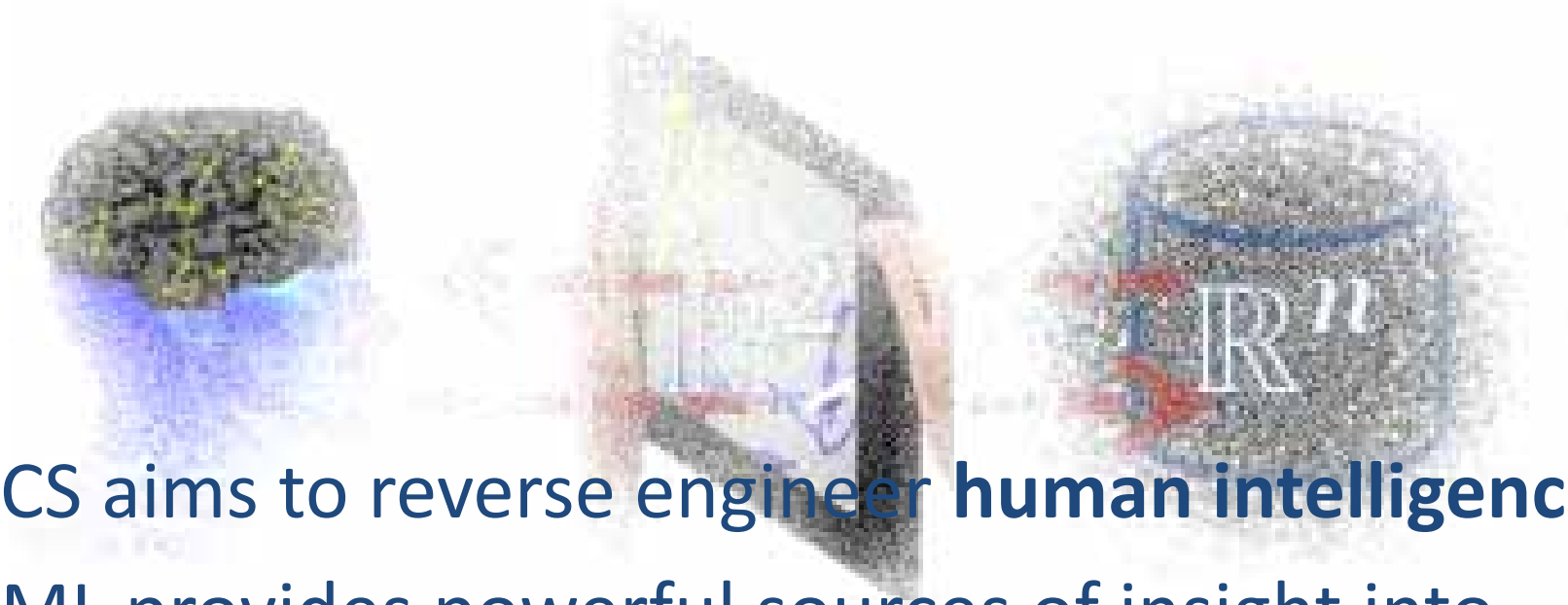
HCI-KDD

Holzinger, A. (2013). Human–Computer Interaction & Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together? In: Lecture Notes in Computer Science 8127 (pp. 319-328)

Cognitive Science vs. Computer Science

- **Cognitive Science → human intelligence**
  - Study the principles of ***human learning*** to understand biological intelligence
- **Human-Computer Interaction → the bridge**
  - Interacting with algorithms that learn shall enhance user friendliness and let concentrate on problem solving - Opening the “black-box” to a “glass-box”
- **Computer Science → computational intelligence**
  - Study the principles of ***machine learning*** to understand artificial intelligence

Why fitting Cognitive Science with Machine Learning?



- CS aims to reverse engineer **human intelligence**;
- ML provides powerful sources of insight into ***how* machine intelligence** can be possible.
- CS therefore raises challenges for, and draws inspiration from ML;
- Insights about the human mind may help inspire **new directions for ML …**

“Solve intelligence – then solve everything else”

https://youtu.be/XAbLn66iHcQ?t=1h28m54s

Demis Hassabis, 22 May 2015

The Royal Society,
Future Directions of Machine Learning Part 2

Google DeepMind

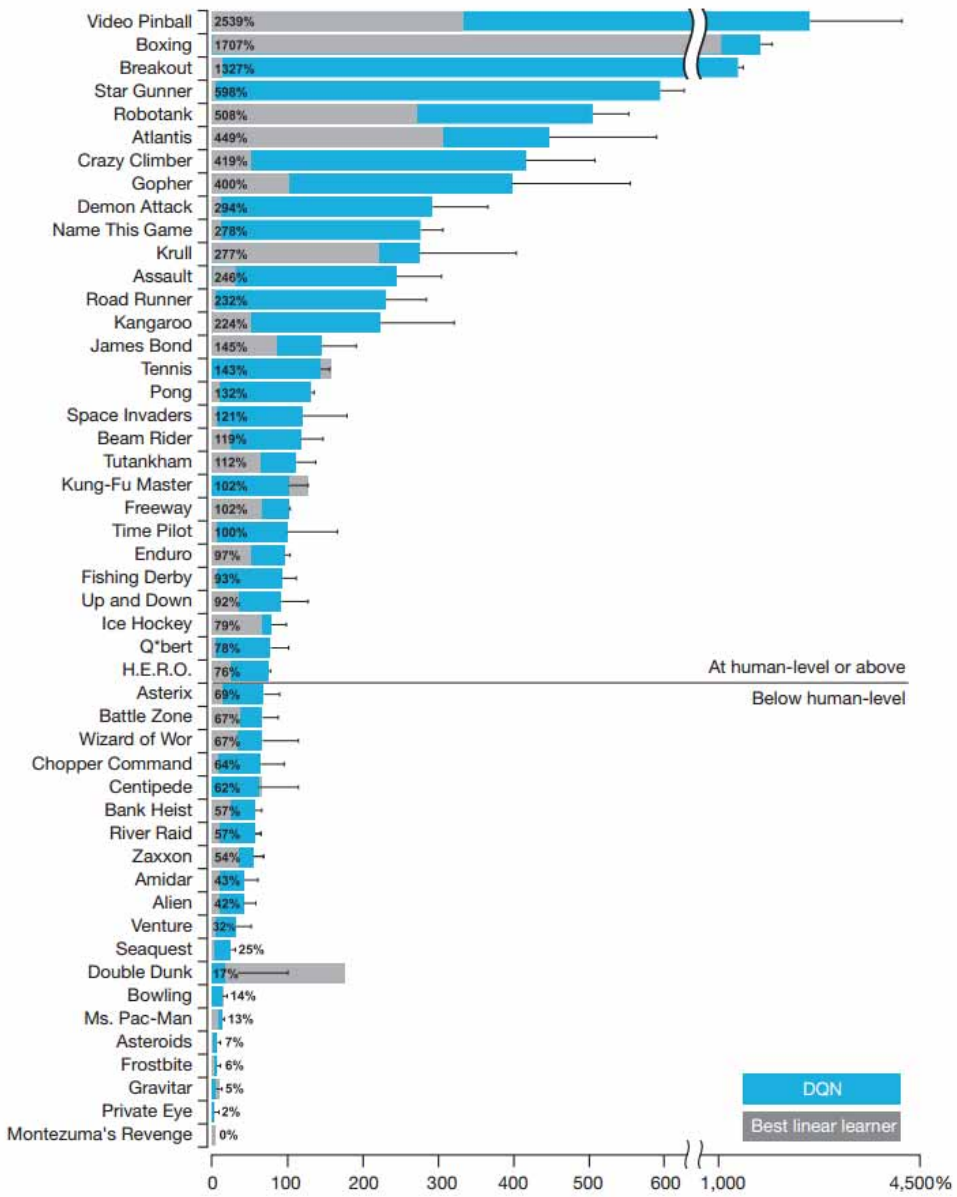
What makes a machine intelligent? Cross-cutting issues

- To hear, to see, to talk, (to smell, taste, touch)
  - Speech recognition, computer vision, natural language processing (olfactory, gustatory sensors)
- To store, to represent, to access
  - Knowledge representation, semantic networks, ontologies, information retrieval
- To reason, to understand, to reflect
  - Logic, Bayesian inference, contextual understanding
- To **learn from data**
  - Improve with experience from previous events

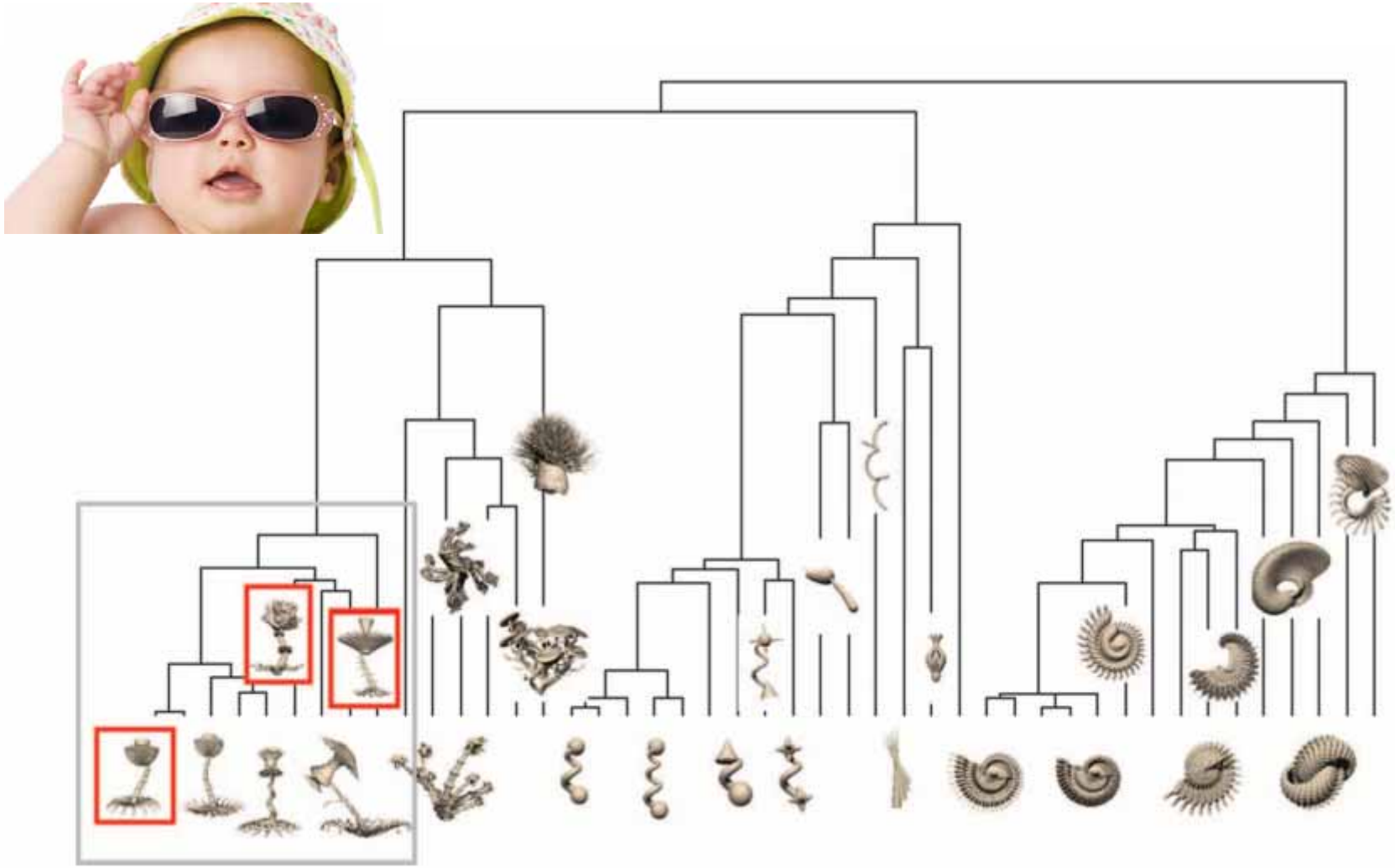
How far are we already ?

Now, compare your best Machine Learning algorithm with a seven year old child …

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. & Hassabis, D. 2015. Human-level control through deep reinforcement learning. Nature, 518, (7540), 529-533, doi:10.1038/nature14236



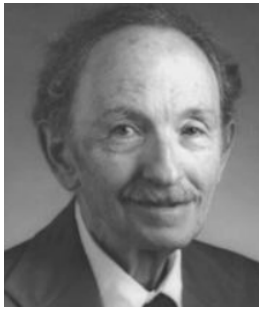
Learning complex concepts from a few examples



Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285, doi:10.1126/science.1192788.

What if we succeed? (Singularity Principle by I.J. Good)

Image credit to http://history.computer.org/pioneers/good.html



"An ultra-intelligent machine could design even better machines; there would then unquestionably be an **"intelligence explosion*"** and the intelligence of man would be left far behind …

It is curious that this point is made so seldom … outside of science fiction."

Irving John Good, Trinity College, Oxford, 1965
Colleague of Alan Turing in Bletchley Park

Good, I. J. 1966. Speculations Concerning the First Ultraintelligent Machine*. In: Franz, L. A. & Morris, R. (eds.) Advances in Computers. Elsevier, pp. 31-88, doi:10.1016/S0065-2458(08)60418-0

https://web.archive.org/web/20010527181244/http://www.aeiveos.com/~bradbury/Authors/Computing/Good-IJ/SCtFUM.html

*) https://intelligence.org/ie-faq/

Today ML is enormously progressing …

- Progress is driven by the explosion in the availability of **big data** and **low-cost computation.**
- **Health is amongst the biggest challenges**

Jordan, M. I. & Mitchell, T. M. 2015. Machine learning: Trends, perspectives, and prospects. Science, 349, (6245), 255-260.

02 Application Area: Health Informatics

Our ML lab … ?

Why is this application area complex ?

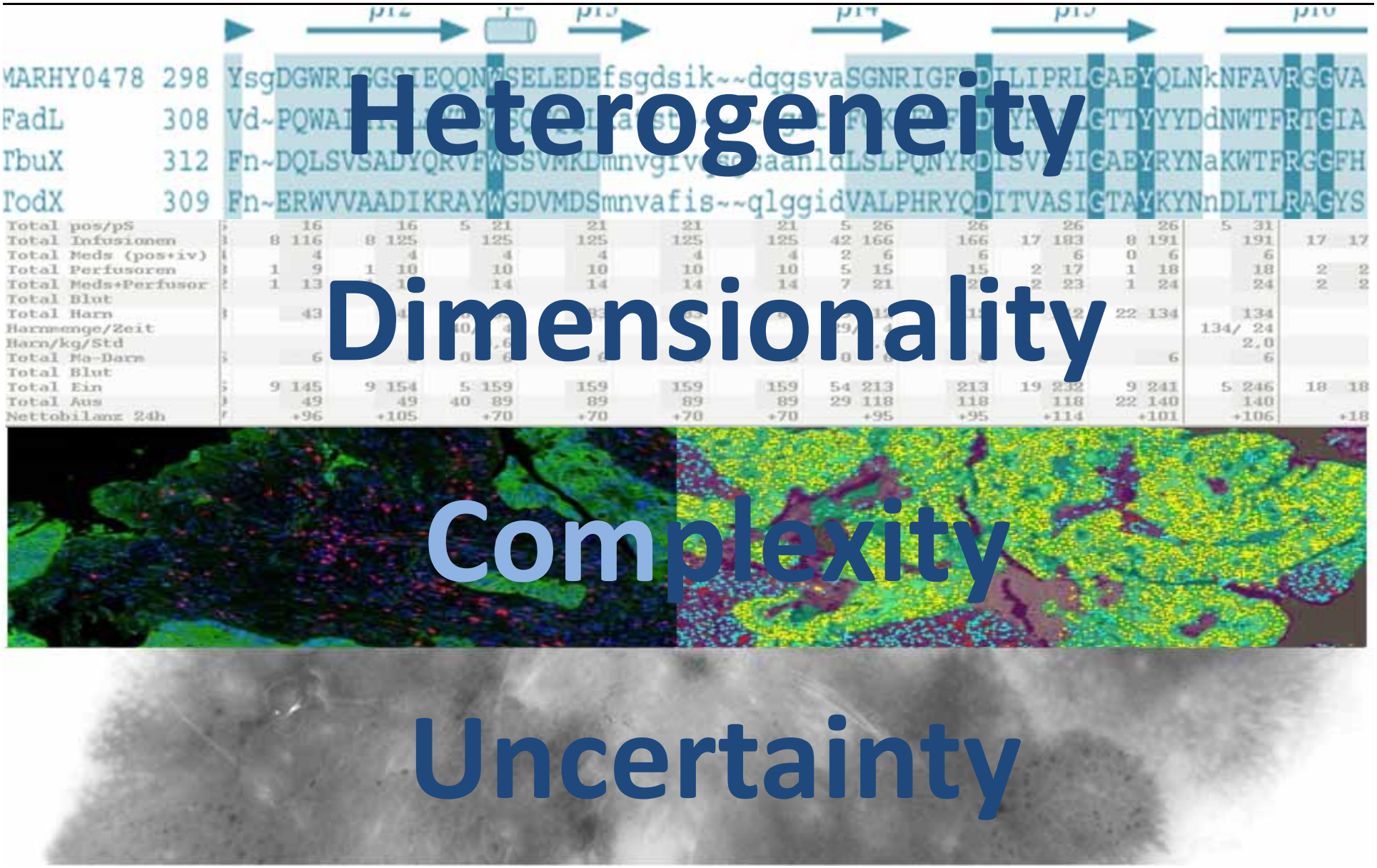
In medicine we have two different worlds …

Our central hypothesis: Information may bridge this gap

Holzinger, A. & Simonic, K.-M. (eds.) 2011. *Information Quality in e-Health. Lecture Notes in Computer Science LNCS 7058, Heidelberg, Berlin, New York: Springer.*

The bridge …

Main problems …



Holzinger, A., Dehmer, M. & Jurisica, I. 2014. Knowledge Discovery and interactive Data Mining in Bioinformatics - State-of-the-Art, future challenges and research directions. BMC Bioinformatics, 15, (S6), I1.

03 Probabilistic Information p(x)

The foundation for machine learning was laid in 1763 …

Thomas Bayes
1701 - 1761

Bayes, T. (1763). An Essay towards solving a Problem in the Doctrine of Chances (Postum communicated by Richard Price). Philosophical Transactions, 53, 370-418.

$$p(x_i) = \sum P(x_i, y_j)$$

$$p(x_i, y_j) = p(y_j|x_i)P(x_i)$$

Bayes' Rule is a corollary of the Sum Rule and Product Rule:

$$p(x_i|y_j) = \frac{p(y_j|x_i)p(x_i)}{\sum p(x_i, y_j)p(x_i)}$$

Barnard, G. A., & Bayes, T. (1958). Studies in the history of probability and statistics: IX. Thomas Bayes's essay towards solving a problem in the doctrine of chances. Biometrika, 45(3/4), 293-315.

d … *data*

h … *hypothesis*

$\mathcal{H}$ … $\{H_1, H_2, \ldots, H_n\}$ $\forall\, h, d$ …

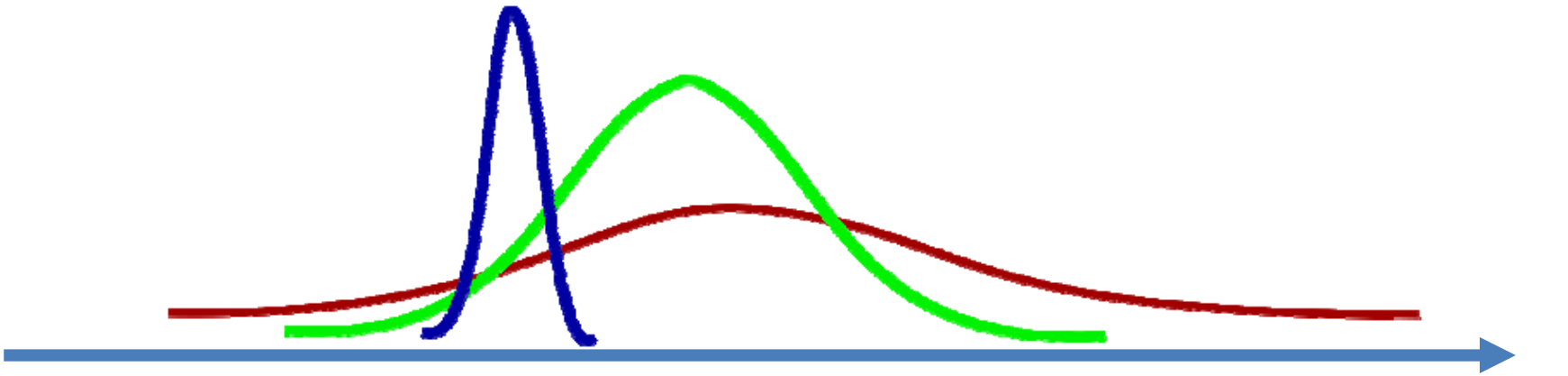
Likelihood

Prior Probability

$$p(h|d) = \frac{p(d|h) * p(h)}{\sum_{h \in H} p(d|h')\, p(h')}$$

Posterior Probability

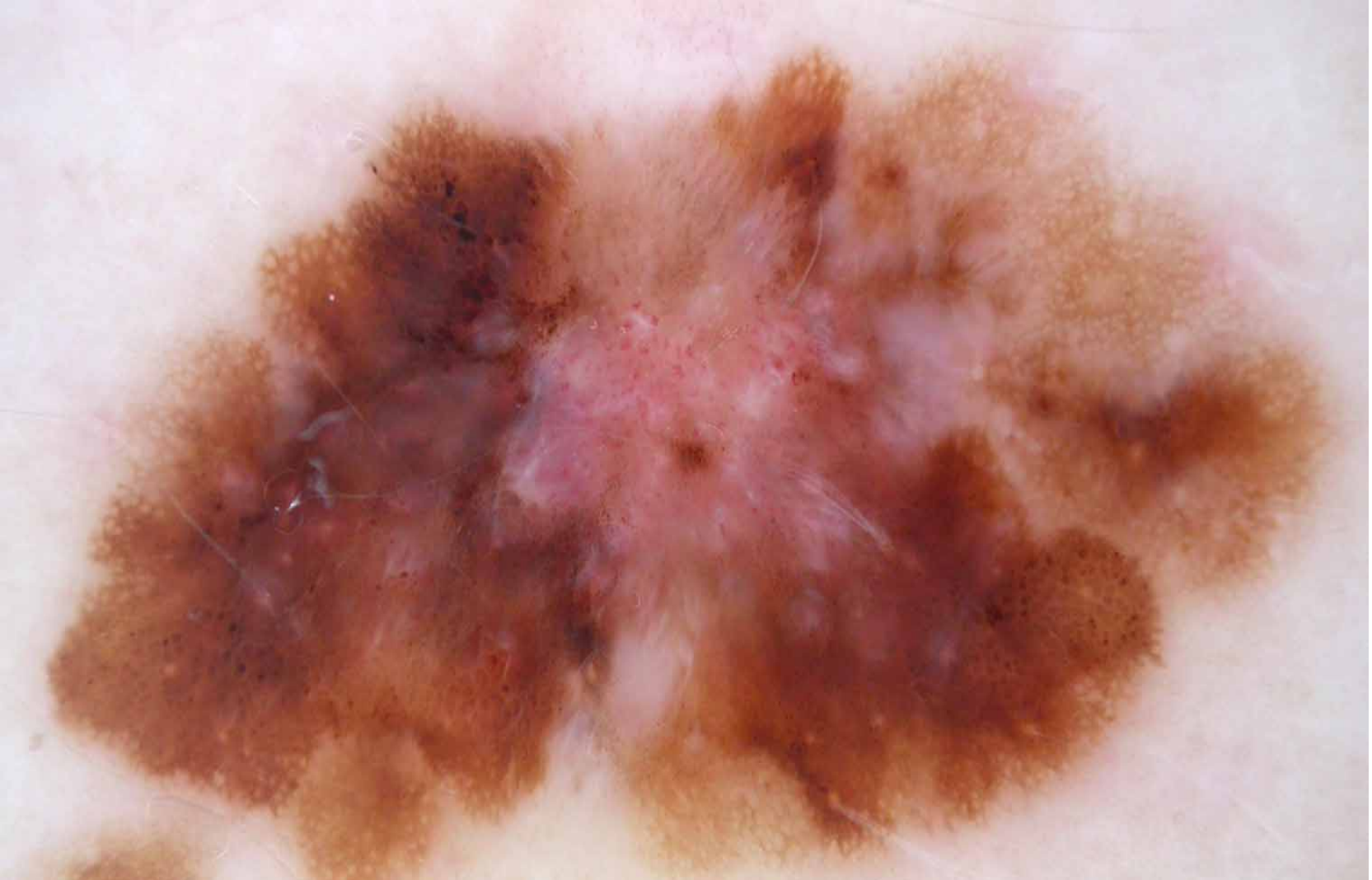
Problem in $\mathbb{R}^n \rightarrow$ complex



Feature parameter θ

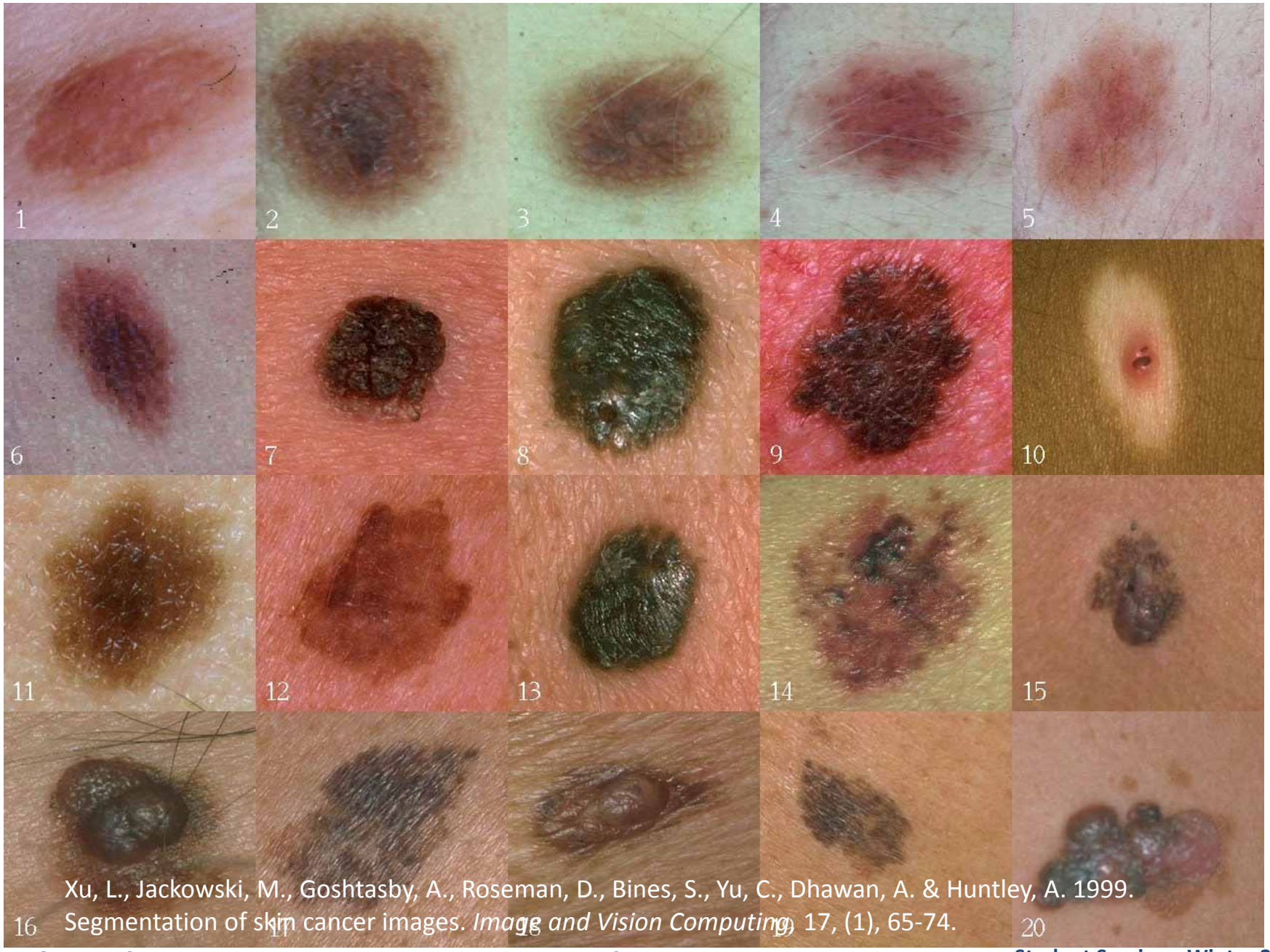
Health Example

To find structural anomalies in such data …



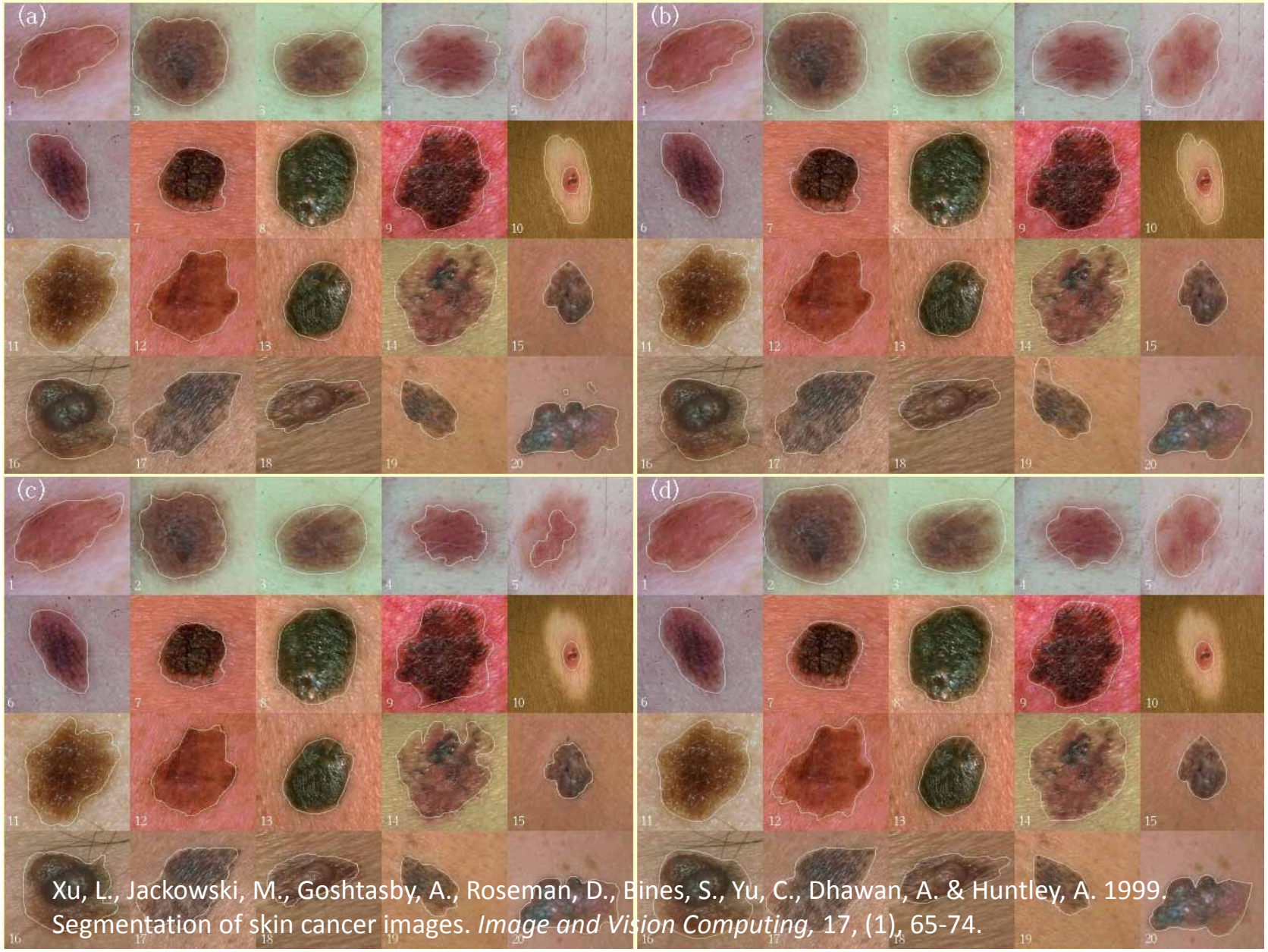
Holzinger, A., Malle, B., Bloice, M., Wiltgen, M., Ferri, M., Stanganelli, I. & Hofmann-Wellenhof, R. 2014. On the Generation of Point Cloud Data Sets: Step One in the Knowledge Discovery Process. *In: Lecture Notes in Computer Science, LNCS 8401.* Berlin Heidelberg: Springer, pp. 57-80, doi:10.1007/978-3-662-43968-5_4.

Learning from previous examples …



Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., Dhawan, A. & Huntley, A. 1999. Segmentation of skin cancer images. *Image and Vision Computing*, 17, (1), 65-74.

The more examples we have the better …



Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., Dhawan, A. & Huntley, A. 1999. Segmentation of skin cancer images. *Image and Vision Computing,* 17, (1), 65-74.

Bayesian Learning from data

$$\mathcal{D} = x_{1:n} = \{x_1, x_2, ..., x_n\}$$

$$p(\mathcal{D}|\theta)$$

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta) * p(\theta)}{p(\mathcal{D})}$$

$$posterior = \frac{likelihood * prior}{evidence}$$

The inverse probability allows to learn from data, infer unknowns, and make predictions

Let the data do the work!

$$\max_{\mathbf{x} \in \mathcal{A} \subset \mathbb{R}^d} f(\mathbf{x})$$

$$p(h|d) \propto p(\mathcal{D}|\theta) * p(h)$$

$$p(f(x)|\mathcal{D}) \propto p(\mathcal{D}|f(x)) * p(f(x)$$

- Machine Learning is the development of algorithms which can **learn from data**
- assessment of **uncertainty,** making **predictions**
- **Automating automation - getting computers to program themselves – let the data do the work!**

Analogies



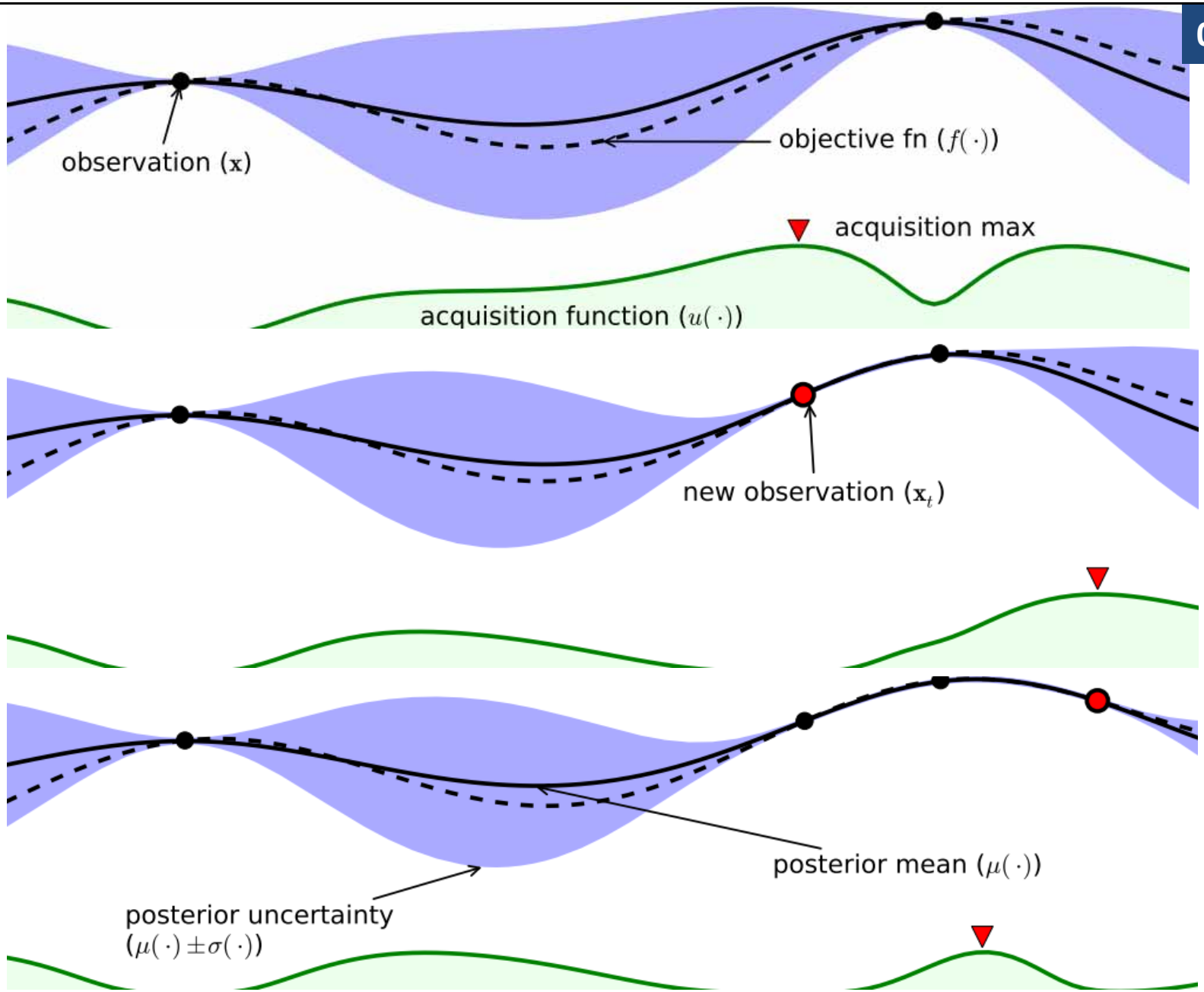

- **Newton, Leibniz, … developed calculus – mathematical language for describing and dealing with rates of change**
- **Bayes, Laplace, … developed probability theory - the mathematical language for describing and dealing with uncertainty**

Gauss integrated those theories

04 Gaussian Processes

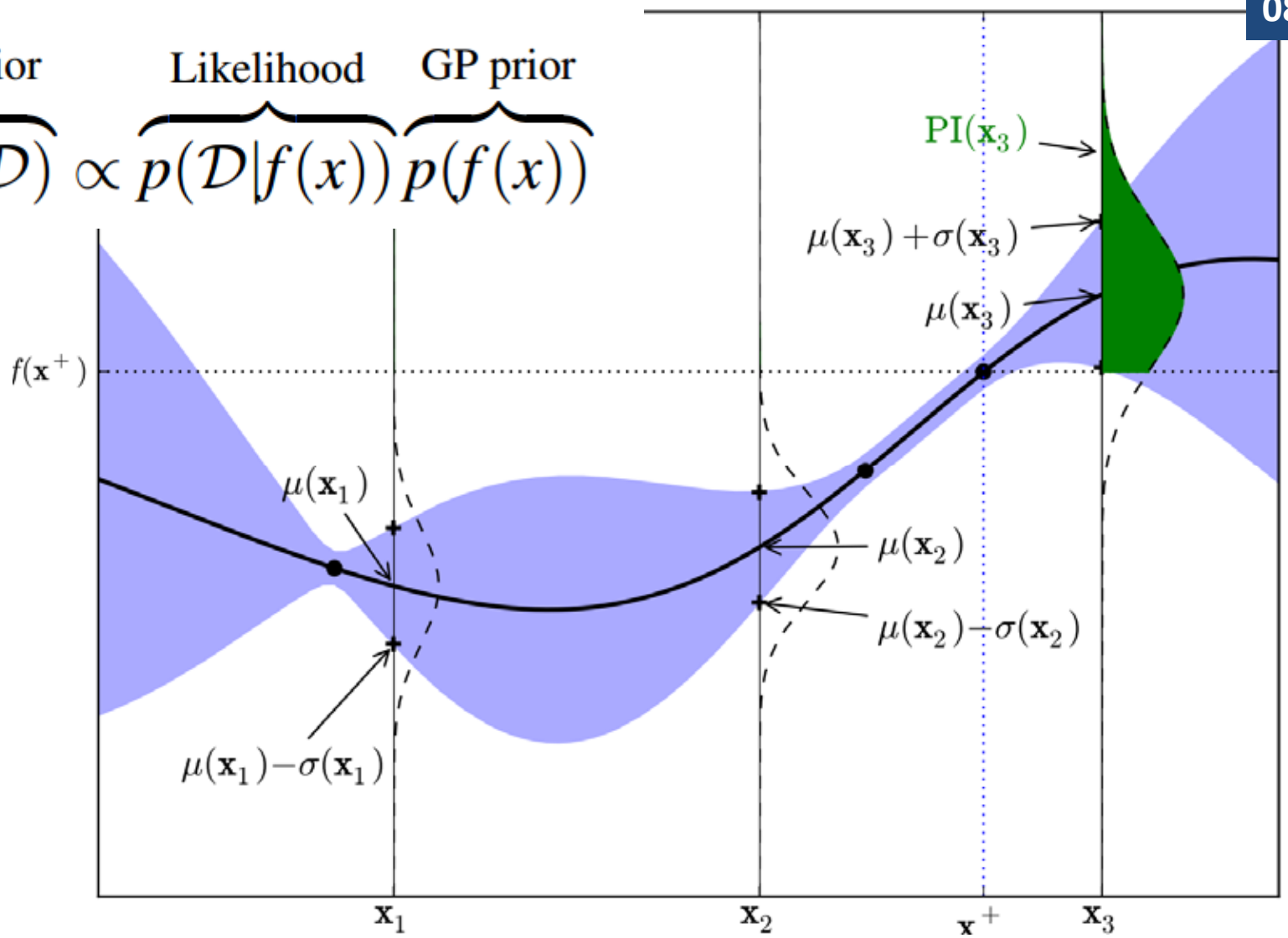
From Bayesian Optimization to Gaussian Process (GP) approximation

Brochu, E., Cora, V. M. & De Freitas, N. 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning arXiv:1012.2599.



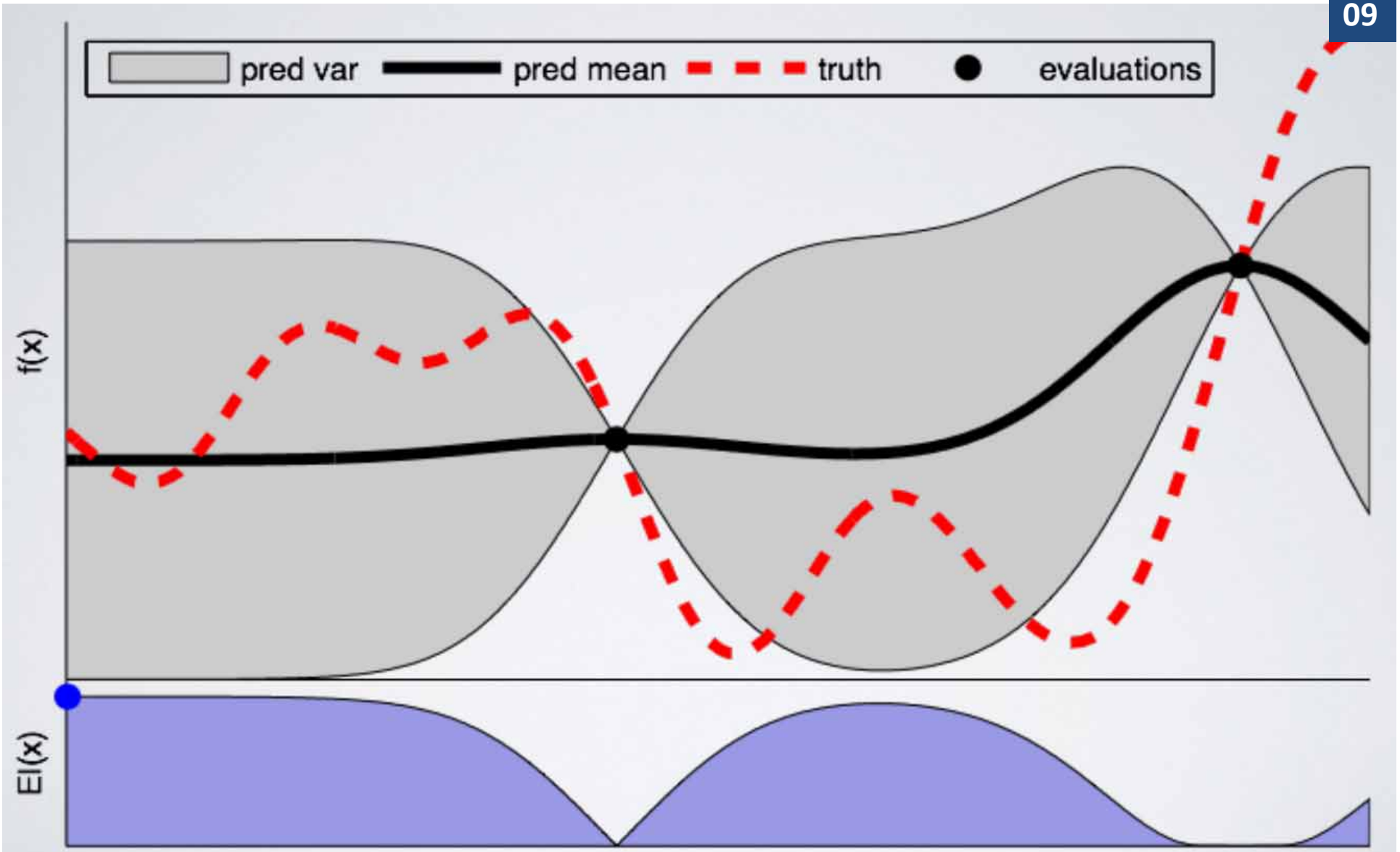
GP = distribution, observations occur in a cont. domain, e.g. t or space

$$\overbrace{p(f(x)|\mathcal{D})}^{\text{GP posterior}} \propto \overbrace{p(\mathcal{D}|f(x))}^{\text{Likelihood}}\overbrace{p(f(x))}^{\text{GP prior}}$$



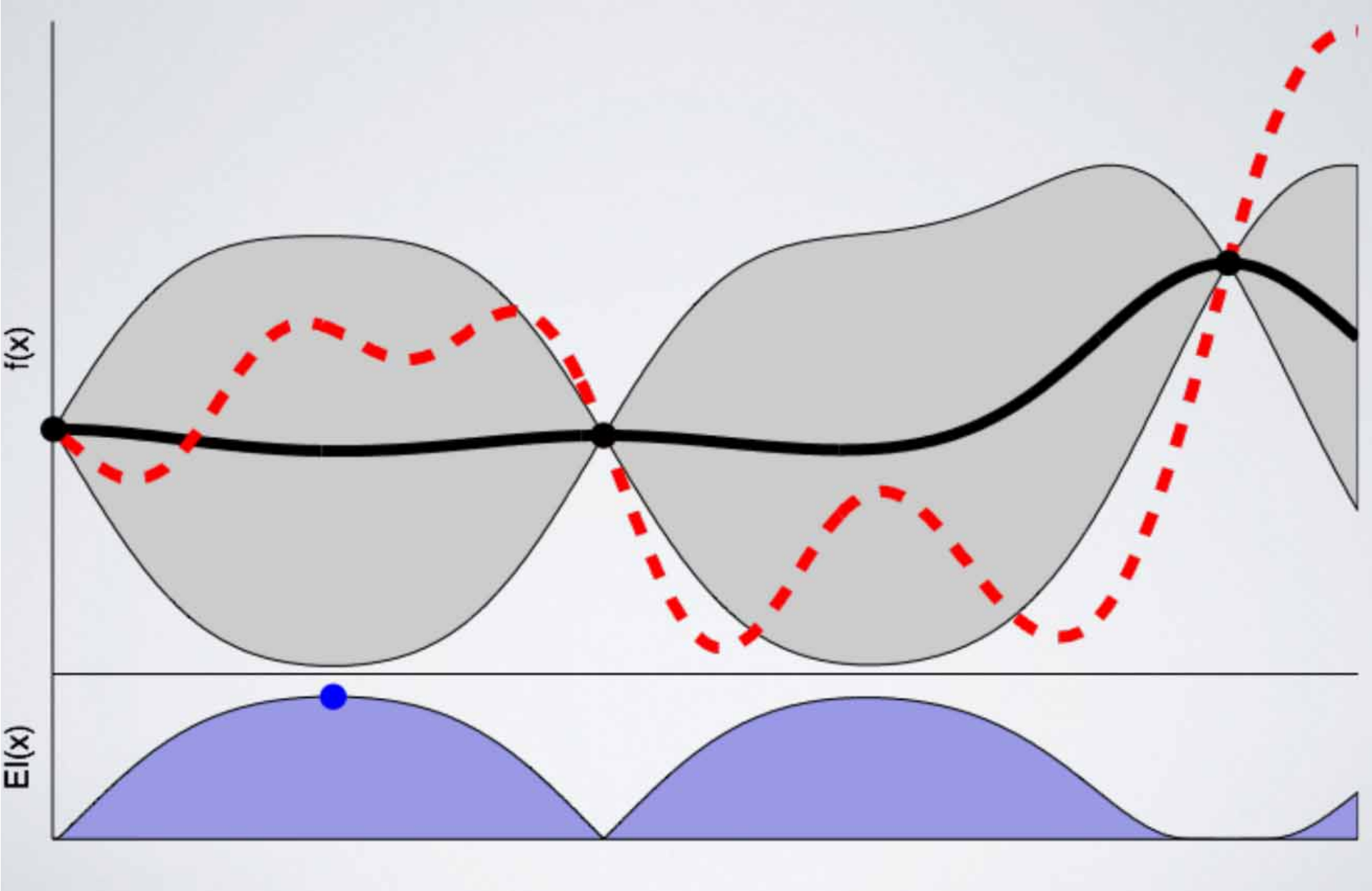
Brochu, E., Cora, V. M. & De Freitas, N. 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv:1012.2599.

Demo on how Bayesian Optimization works …

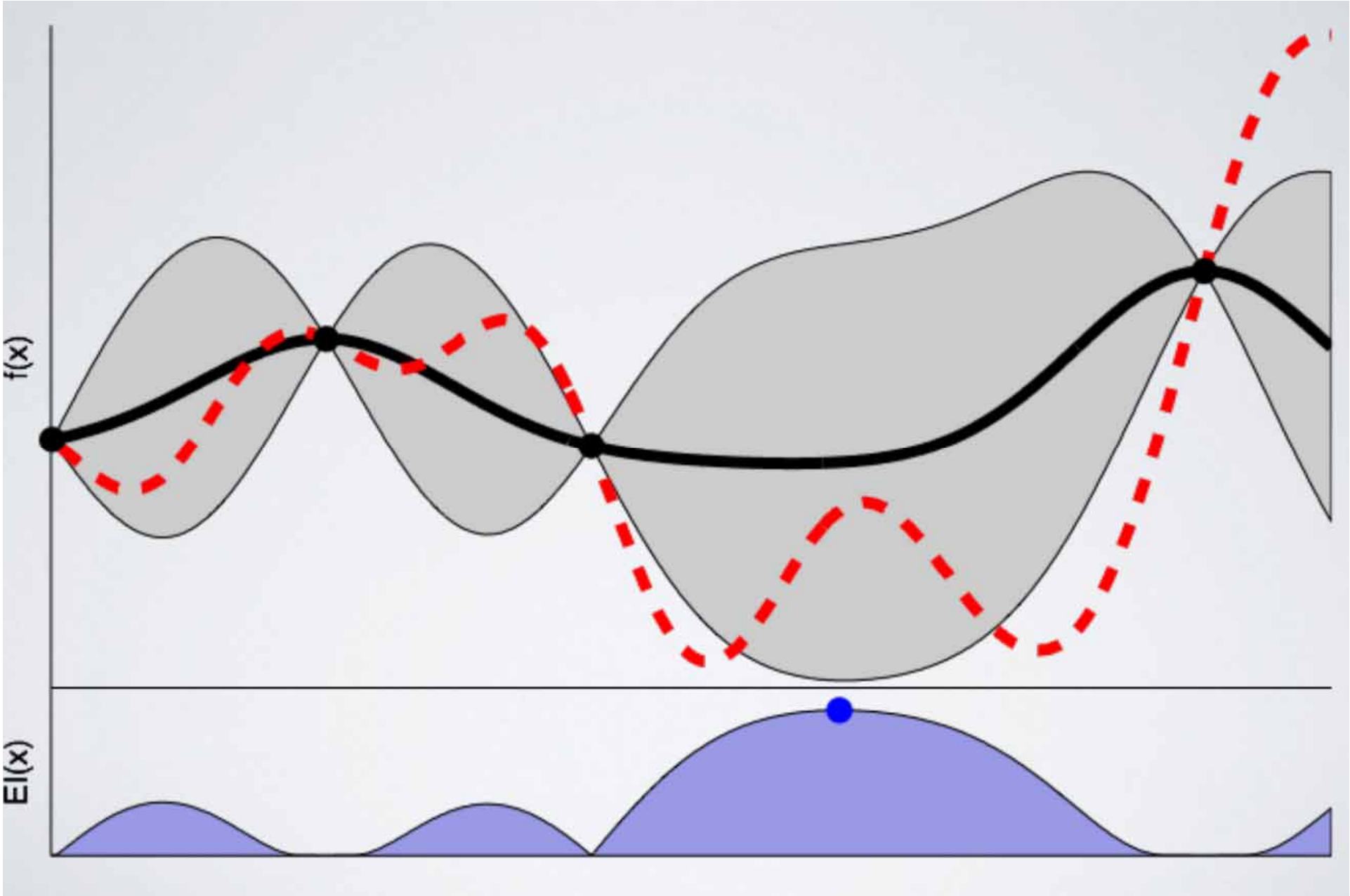


Snoek, J., Larochelle, H. & Adams, R. P. Practical bayesian optimization of machine learning algorithms. Advances in neural information processing systems, 2012. 2951-2959.

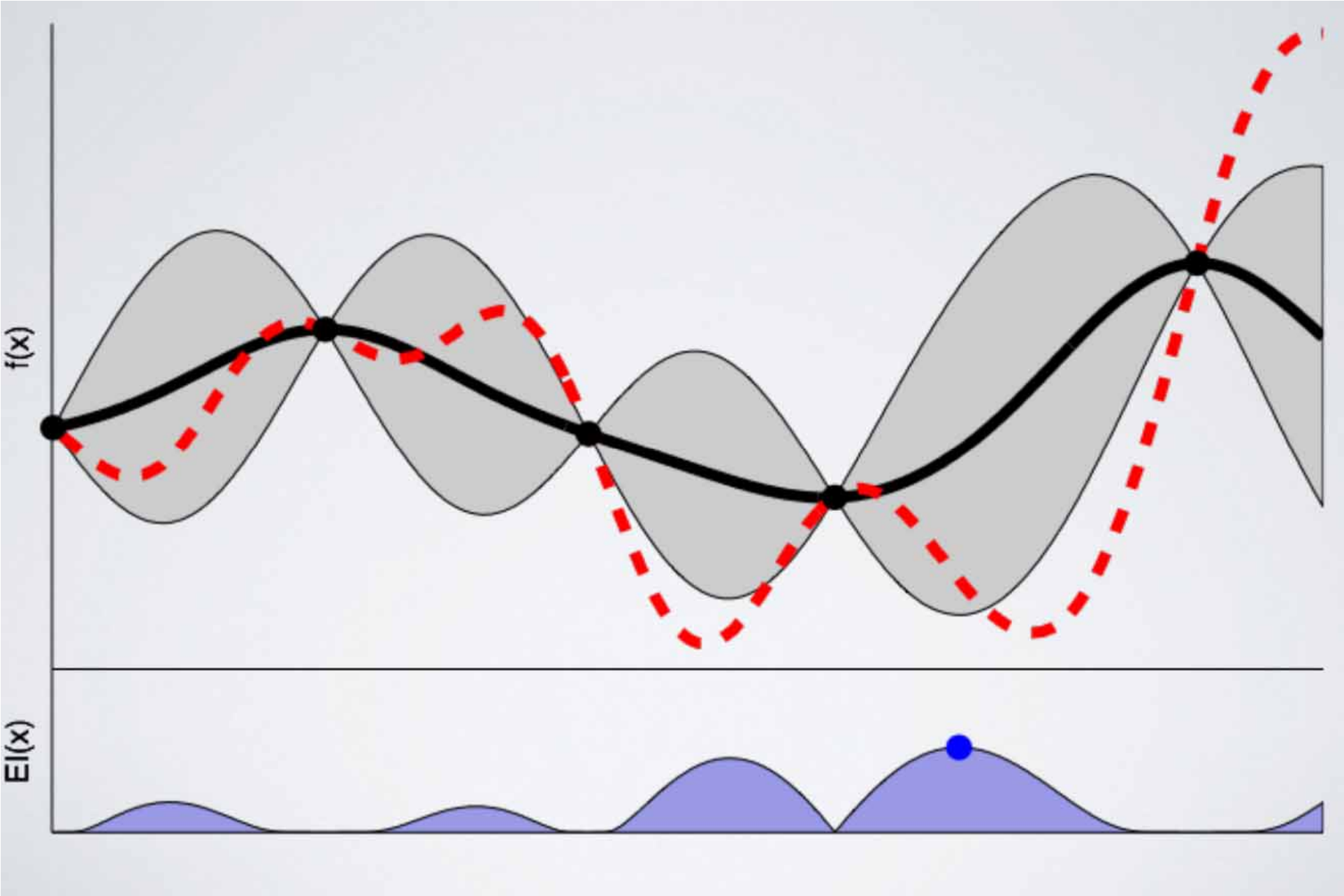
Bayesian Optimization 1



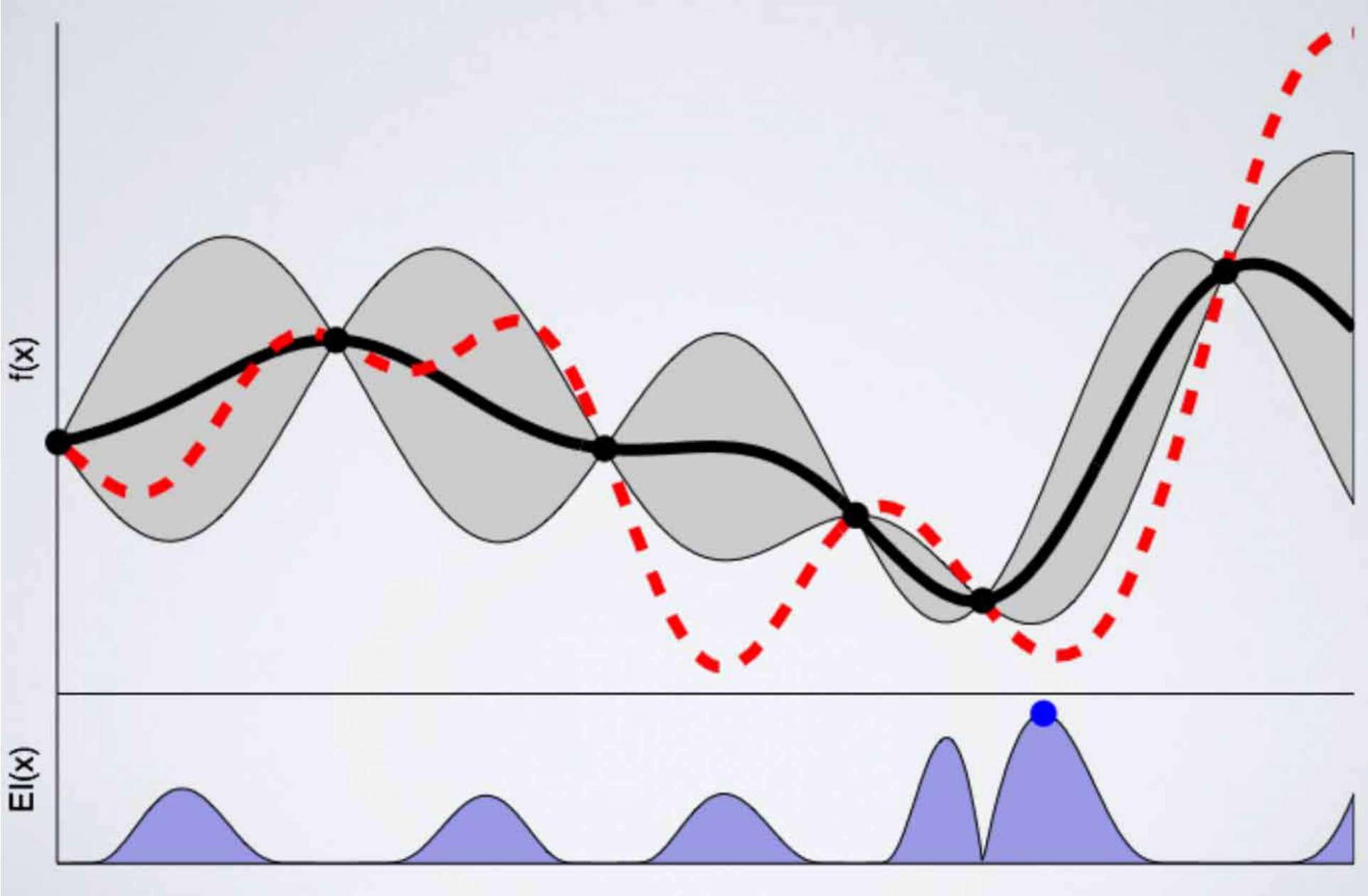
Bayesian Optimization 2



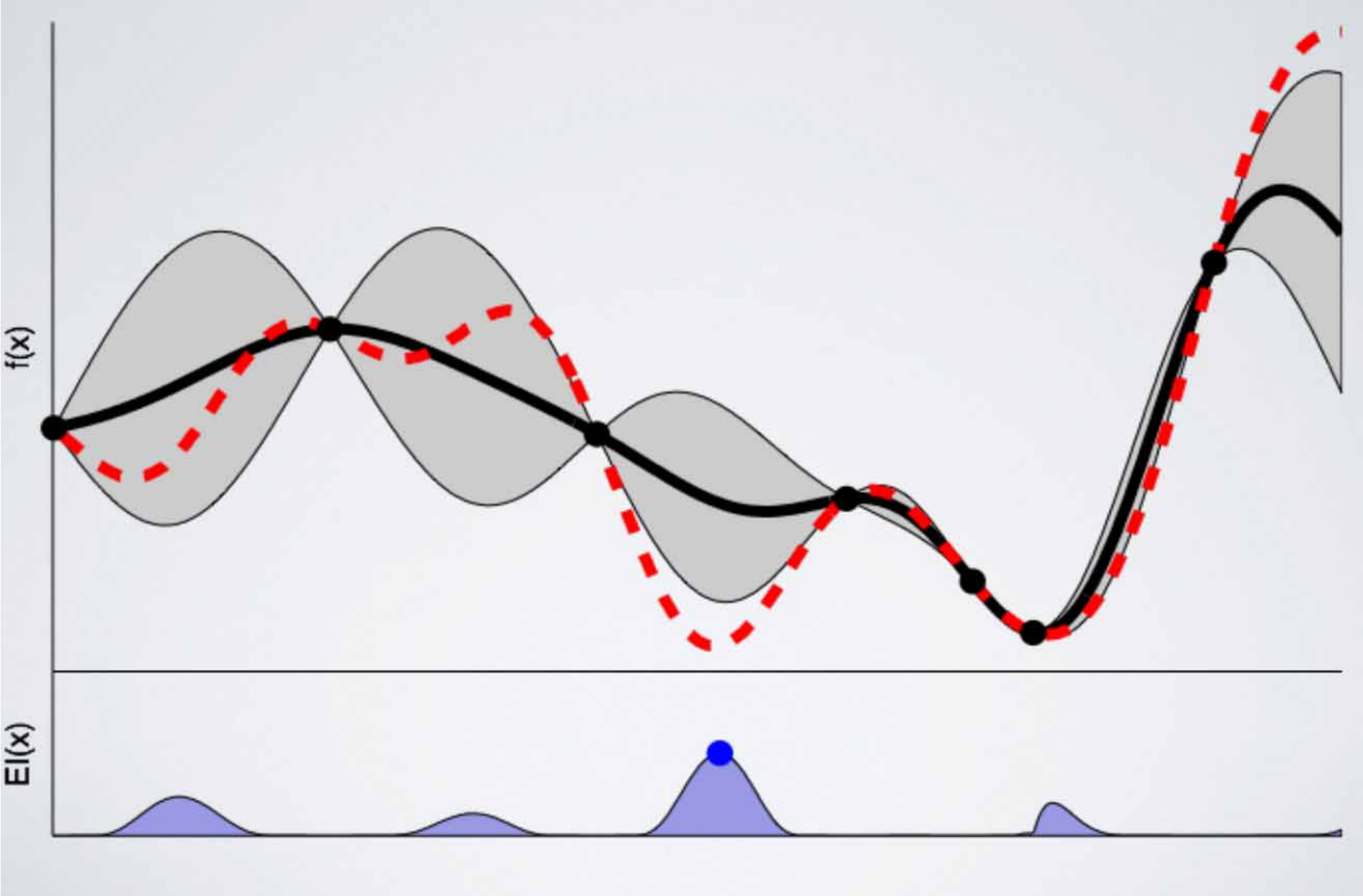
Bayesian Optimization 3



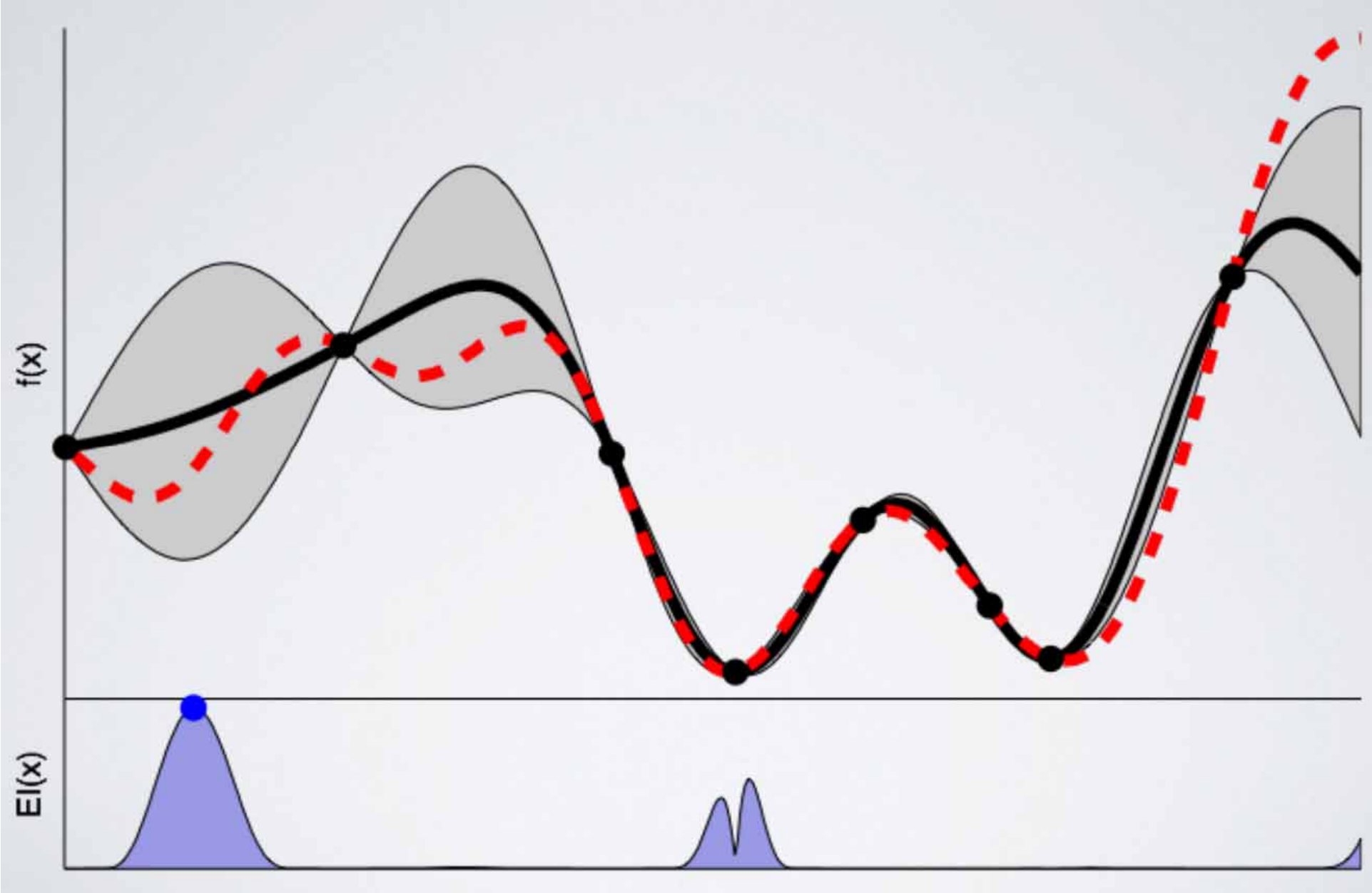
Bayesian Optimization 4



Bayesian Optimization 5



Bayesian Optimization 6



Why is this relevant for health informatics?

Reasoning under uncertainty

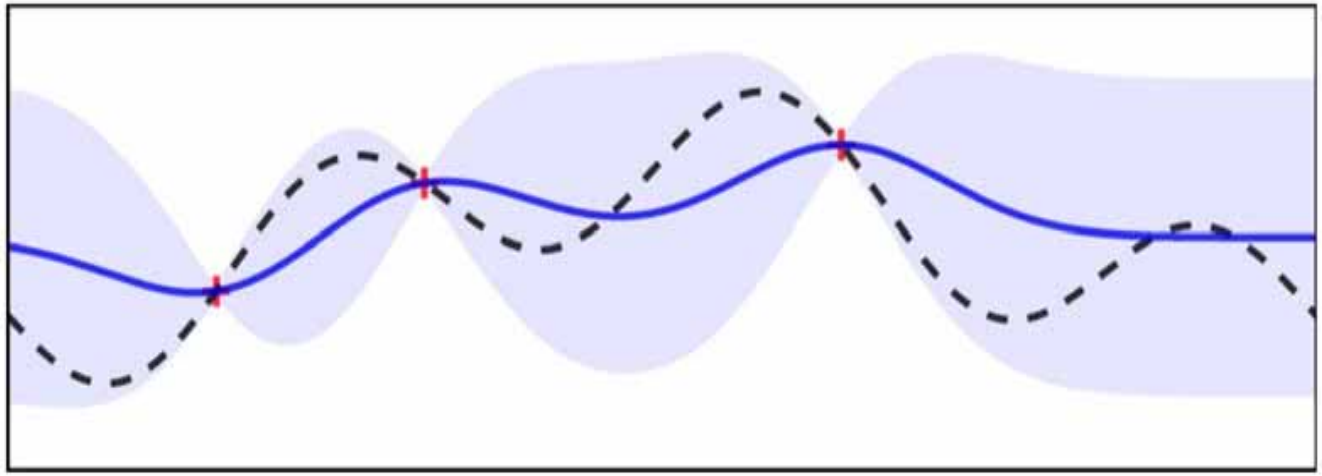
- Take patient information, e.g., observations, symptoms, test results, -omics data, etc. etc.
- Reach conclusions, and **predict** into the future, e.g. how likely will a patient be readmissioned
- Prior = belief before making a particular observation
- Posterior = belief after making the observation and is the prior for the next observation – intrinsically incremental

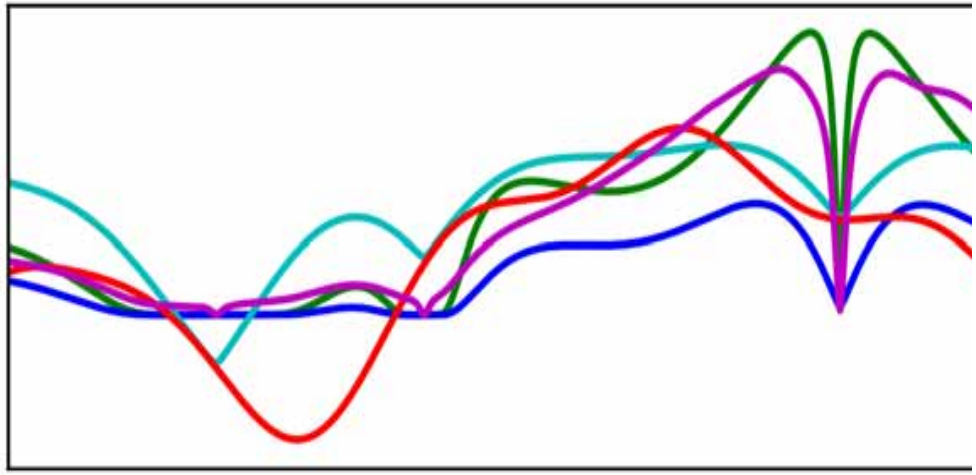
$$p(x_i|y_j) = \frac{p(y_j|x_i)p(x_i)}{\sum p(x_i, y_j)p(x_i)}$$

… big data is good for automatic Machine Learning

and the grand goal of aML is …

taking the human out-of-the-loop





Algorithm 1: Bayesian optimization

1: **for** $n = 1, 2, \ldots,$ **do**
2: select new $\mathbf{x}_{n+1}$ by optimizing acquisition function α

$$\mathbf{x}_{n+1} = \arg\max_{\mathbf{x}} \alpha(\mathbf{x}; \mathcal{D}_n)$$

3: query objective function to obtain y_{n+1}
4: augment data $\mathcal{D}_{n+1} = \{\mathcal{D}_n, (\mathbf{x}_{n+1}, y_{n+1})\}$
5: update statistical model
6: **end for**

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & De Freitas, N. 2016. **Taking the human out of the loop:** A review of Bayesian optimization. *Proceedings of the IEEE,* 104, (1), 148-175, doi:10.1109/JPROC.2015.2494218.

05 aML

Everything is machine learning …

- Today most ML-applications are using automatic Machine Learning (aML) approaches

- automatic Machine Learning (aML) := algorithms which interact with agents and can optimize their learning behaviour trough this interaction

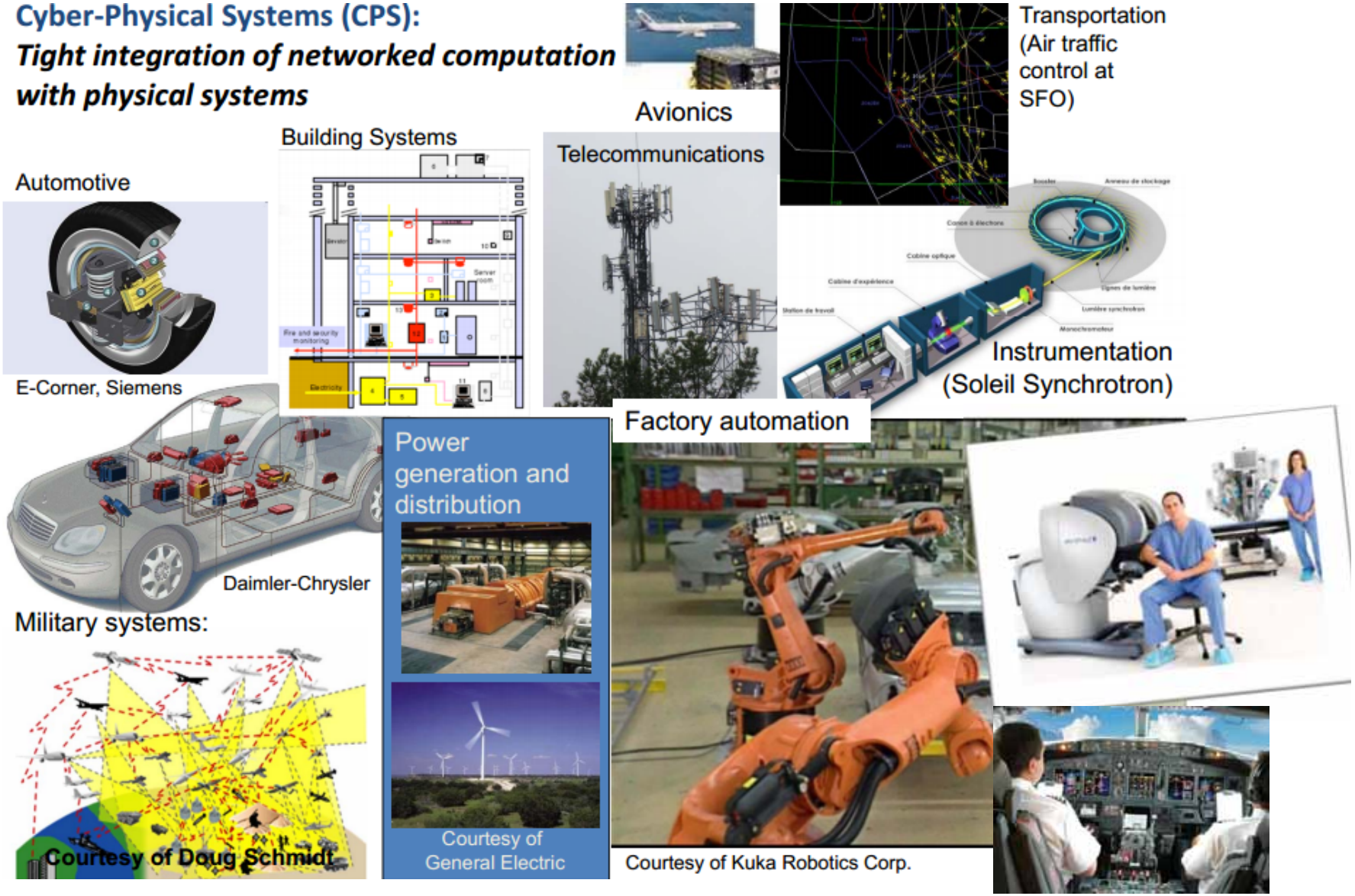
Jordan, M. I. & Mitchell, T. M. 2015. Machine learning: Trends, perspectives, and prospects. Science, 349, (6245), 255-260, doi:10.1126/science.aaa8415.

Best practice examples of aML

Fully automatic autonomous vehicles ("Google car")

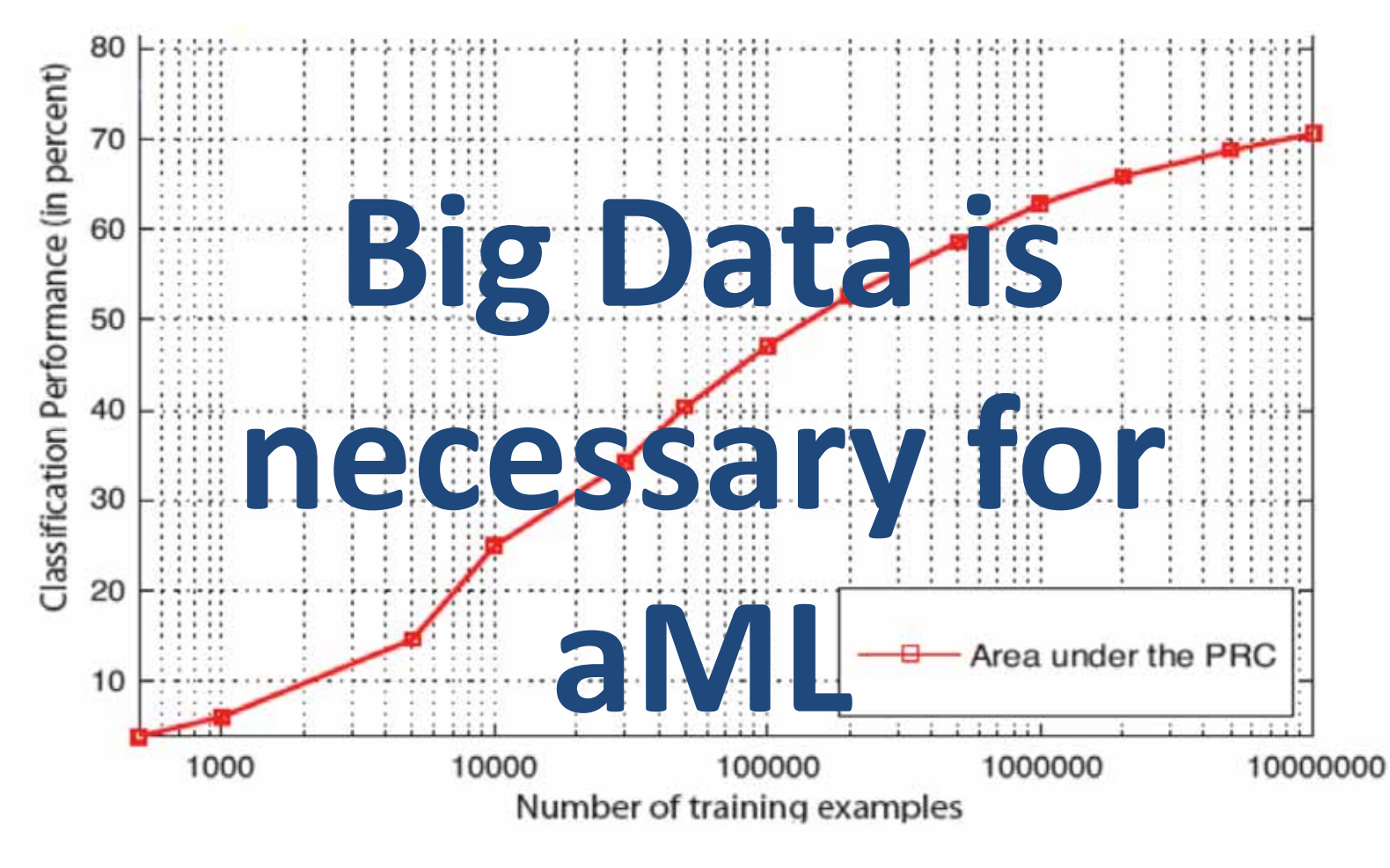
Dietterich, T. G. & Horvitz, E. J. 2015. Rise of concerns about AI: reflections and directions. Communications of the ACM, 58, (10), 38-40.

… and thousands of industrial aML applications …



Seshia, S. A., Juniwal, G., Sadigh, D., Donze, A., Li, W., Jensen, J. C., Jin, X., Deshmukh, J., Lee, E. & Sastry, S. 2015. Verification by, for, and of Humans: Formal Methods for Cyber-Physical Systems and Beyond. Illinois ECE Colloquium.

Big Data is necessary for aML !



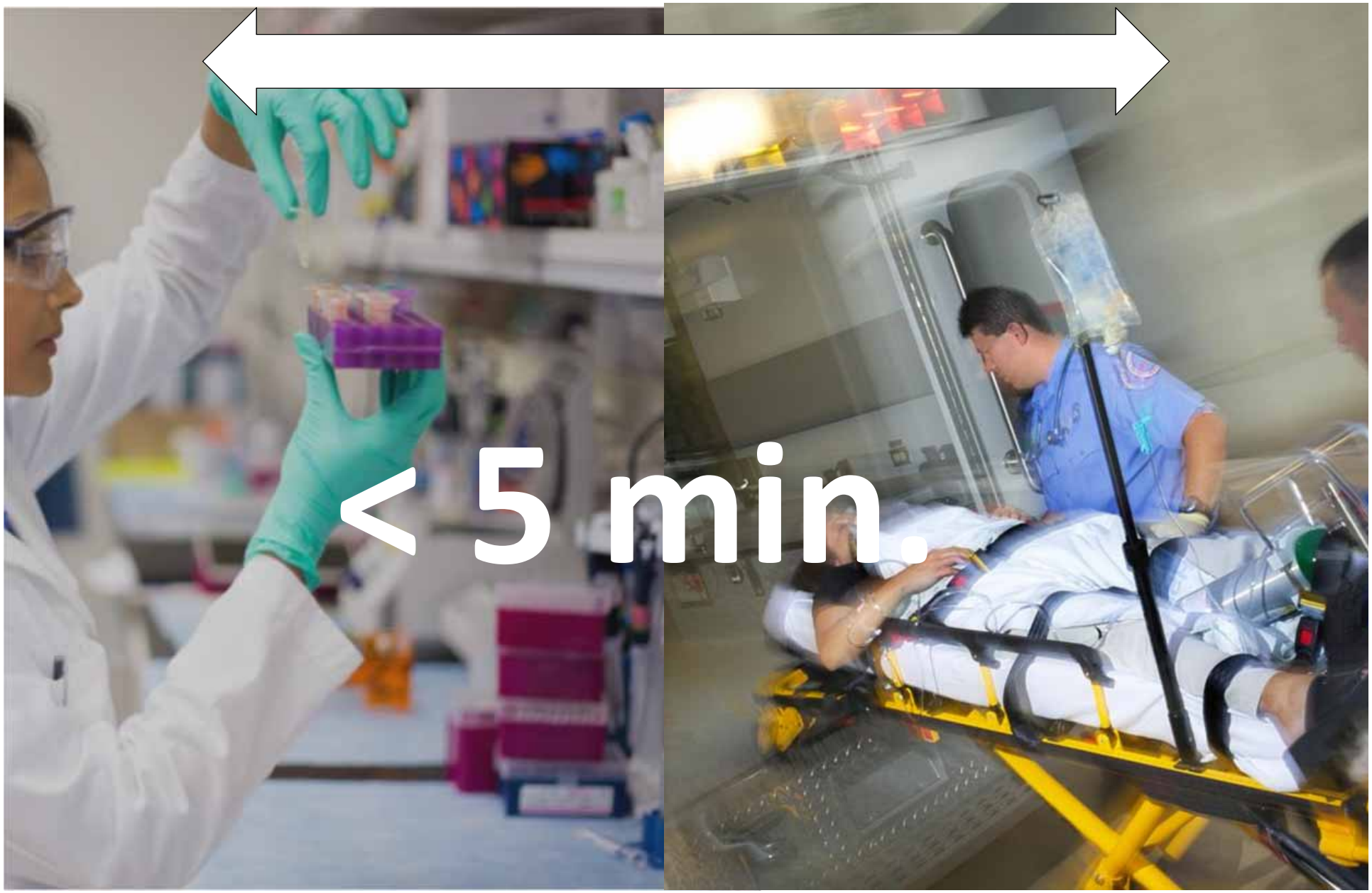
Big Data is necessary for aML

Sonnenburg, S., Rätsch, G., Schäfer, C. & Schölkopf, B. 2006. Large scale multiple kernel learning. Journal of Machine Learning Research, 7, (7), 1531-1565.

Does this all work here as well?

Medical Decision Making as a Search Task in $\mathcal{H}$ Problem: Time (t)

Search in an arbitrarily high-dimensional space < 5 min.!



06 iML

When does aML fail …

- Sometimes we do **not** have "big data", where aML-algorithms benefit.
- Sometimes we have
  - **Small amount of data sets**
  - Rare Events – **no training samples**
  - **NP-hard problems, e.g.**
    - Subspace Clustering,
    - Protein-Folding,
    - k-Anonymization,
    - Category Discovery, etc. etc....

Sometimes we (still) need a human-in-the-loop

Sometimes we need a doctor-in-the-loop

Definition of iML

A6

- interactive Machine Learning (iML) := algorithms which interact with agents*) and can optimize their learning behaviour trough this interaction
- ***)where the agents can be human**

Holzinger, A. 2015. Interactive Machine Learning (iML). Informatik Spektrum DOI: 10.1007/s00287-015-0941-6

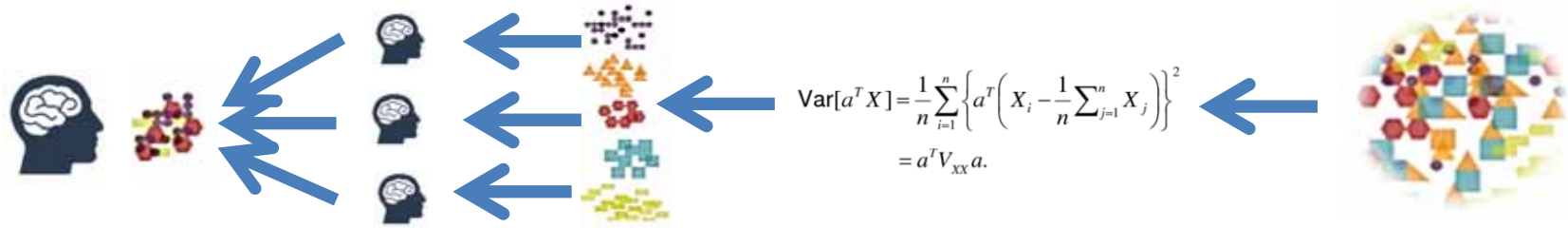
Holzinger, A. 2016. Interactive Machine Learning for Health Informatics: When do we need the human-in-the-loop? Springer Brain Informatics (BRIN), 3, (2), 119-131, doi:10.1007/s40708-016-0042-6.

A group of experts-in-the-loop

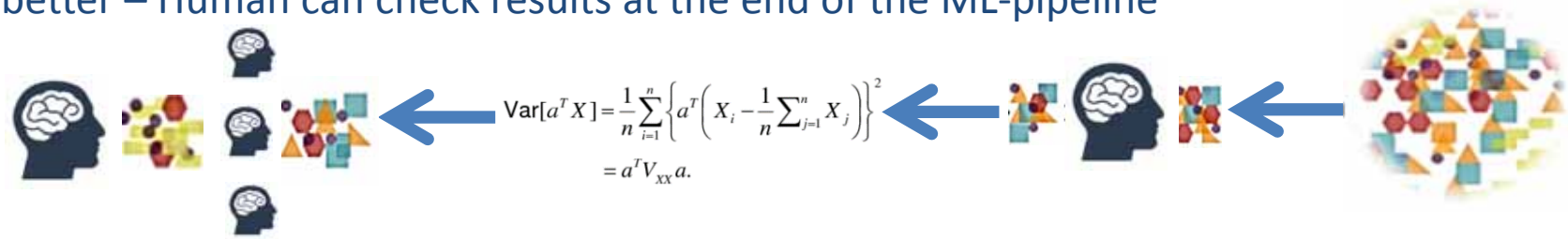
A crowd of people-in-the-loop

aML: Unsupervised – Supervised – Semi-supervised

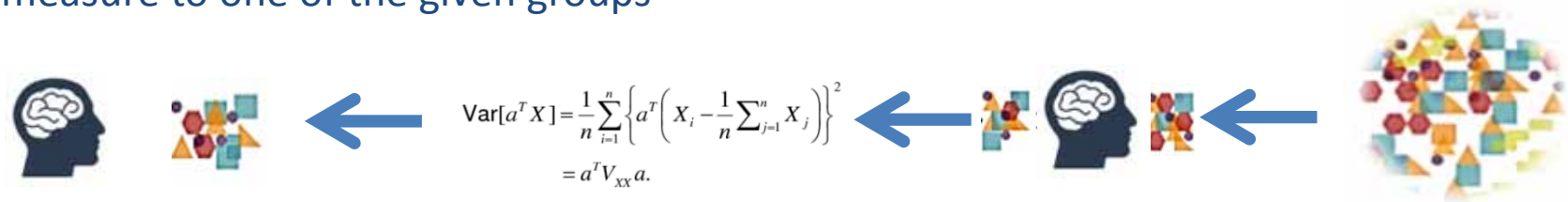
A) Unsupervised ML: Algorithm is applied on the raw data and learns fully automatic – Human can check results at the end of the ML-pipeline



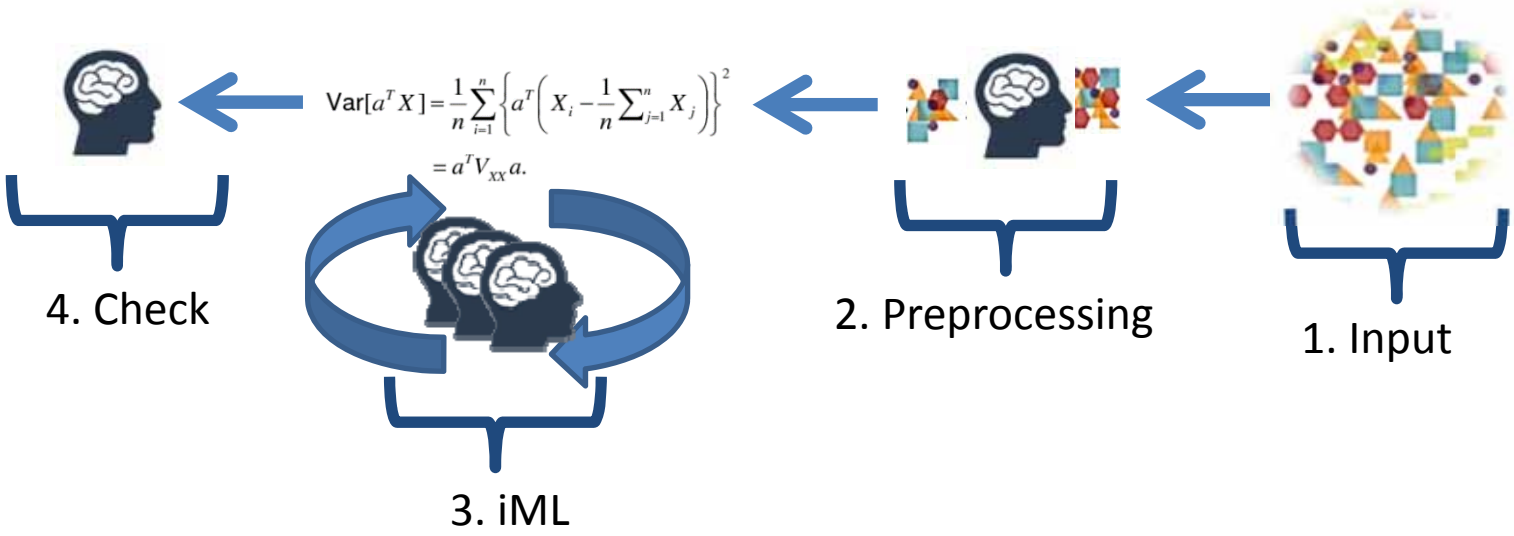
B) Supervised ML: Humans are providing the labels for the training data and/or select features to feed the algorithm to learn – the more samples the better – Human can check results at the end of the ML-pipeline



C) Semi-Supervised Machine Learning: A mixture of A and B – mixing labeled and unlabeled data so that the algorithm can find labels according to a similarity measure to one of the given groups



D) **Interactive Machine Learning:** Human is seen as an agent involved in the actual learning phase, step-by-step influencing measures such as distance, cost functions …



Constraints of humans: Robustness, subjectivity, transfer?
Open Questions: Evaluation, replicability, …

Holzinger, A. 2016. Interactive Machine Learning for Health Informatics: When do we need the human-in-the-loop? Brain Informatics (BRIN), 3, (2), 119-131, doi:10.1007/s40708-016-0042-6.

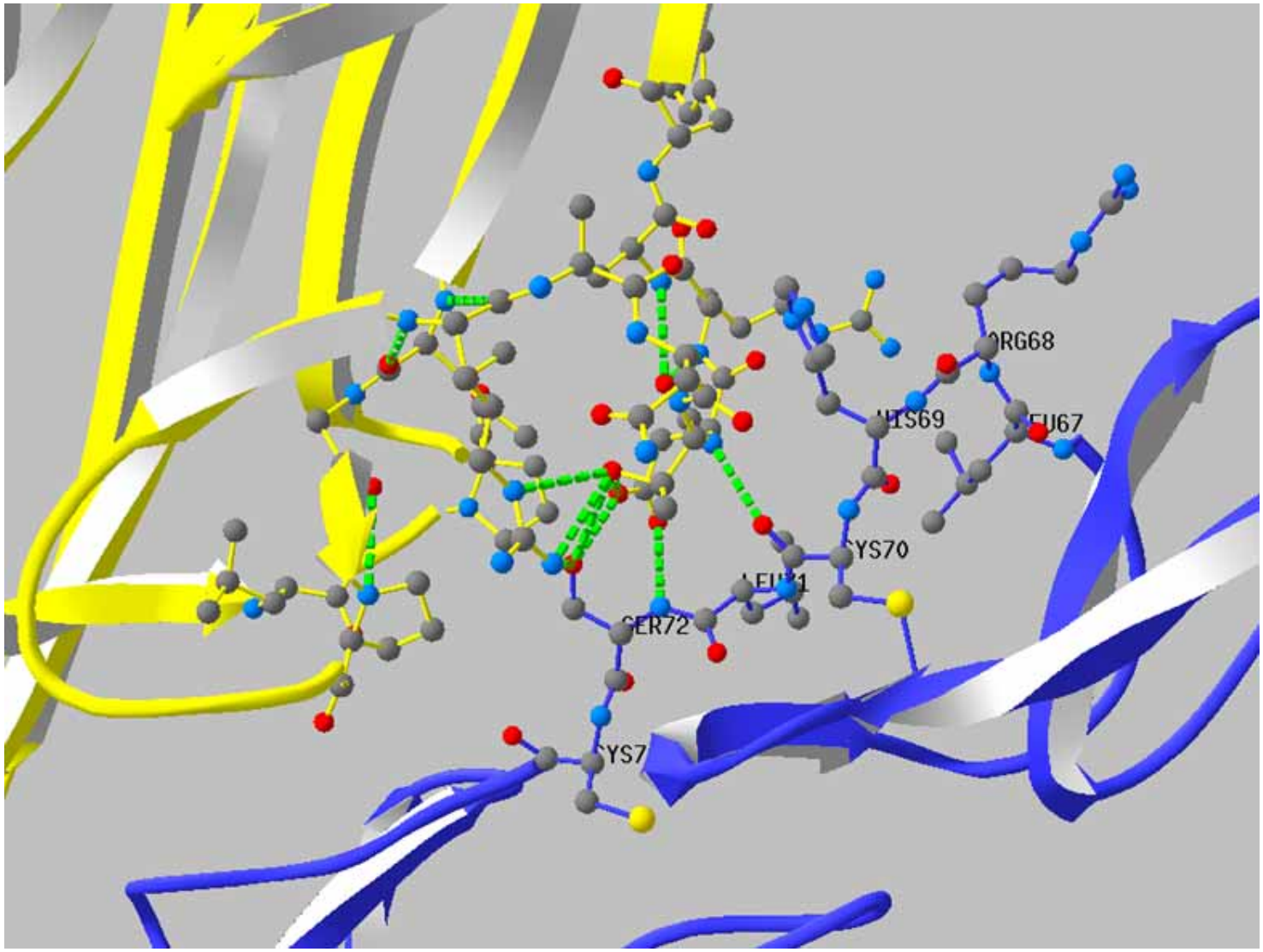
Three examples for the usefulness of the iML approach

- Example 1: Subspace Clustering [1]
- Example 2: k-Anonymization [2]
- Example 3: Protein Design [3]

[1] Hund, M., Böhm, D., Sturm, W., Sedlmair, M., Schreck, T., Ullrich, T., Keim, D. A., Majnaric, L. & Holzinger, A. 2016. Visual analytics for concept exploration in subspaces of patient groups: Making sense of complex datasets with the Doctor-in-the-loop. Brain Informatics, 1-15, doi:10.1007/s40708-016-0043-5.
[2] Kieseberg, P., Malle, B., Frühwirt, P., Weippl, E. & Holzinger, A. 2016. A tamper-proof audit and control system for the doctor in the loop. Brain Informatics, 1-11, doi:10.1007/s40708-016-0046-2.
[3] Jeanquartier, F., Jean-Quartier, C. & Holzinger, A. 2015. Integrated web visualizations for protein-protein interaction databases. BMC Bioinformatics, 16, (1), 195, doi:10.1186/s12859-015-0615-z.

Example 3: Proteins are the building blocks of life …

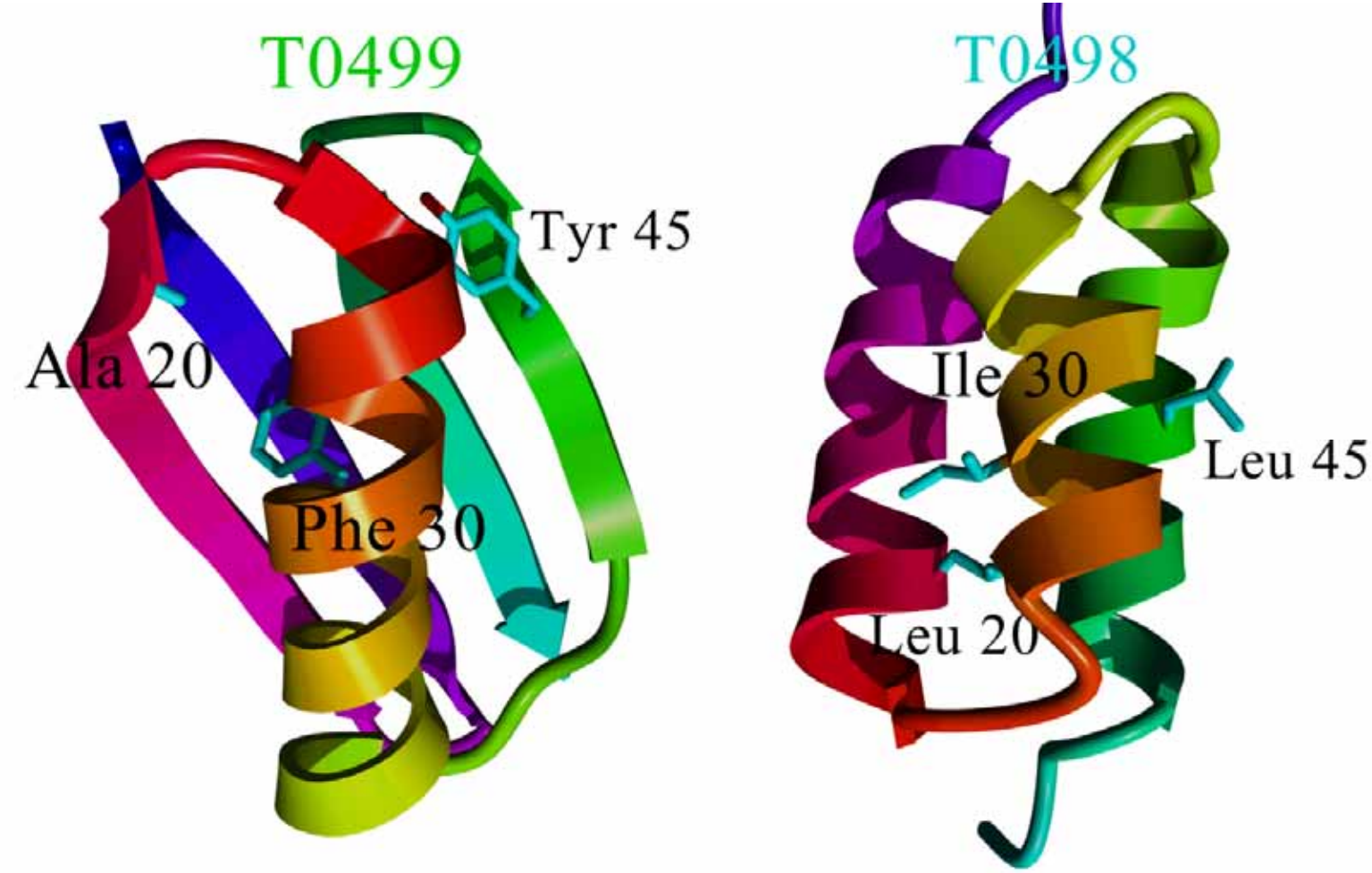
A4



Wiltgen, M., Holzinger, A. & Tilz, G. P. (2007) Interactive Analysis and Visualization of Macromolecular Interfaces Between Proteins. In: *Lecture Notes in Computer Science (LNCS 4799). Berlin, Heidelberg, New York, Springer, 199-212.*

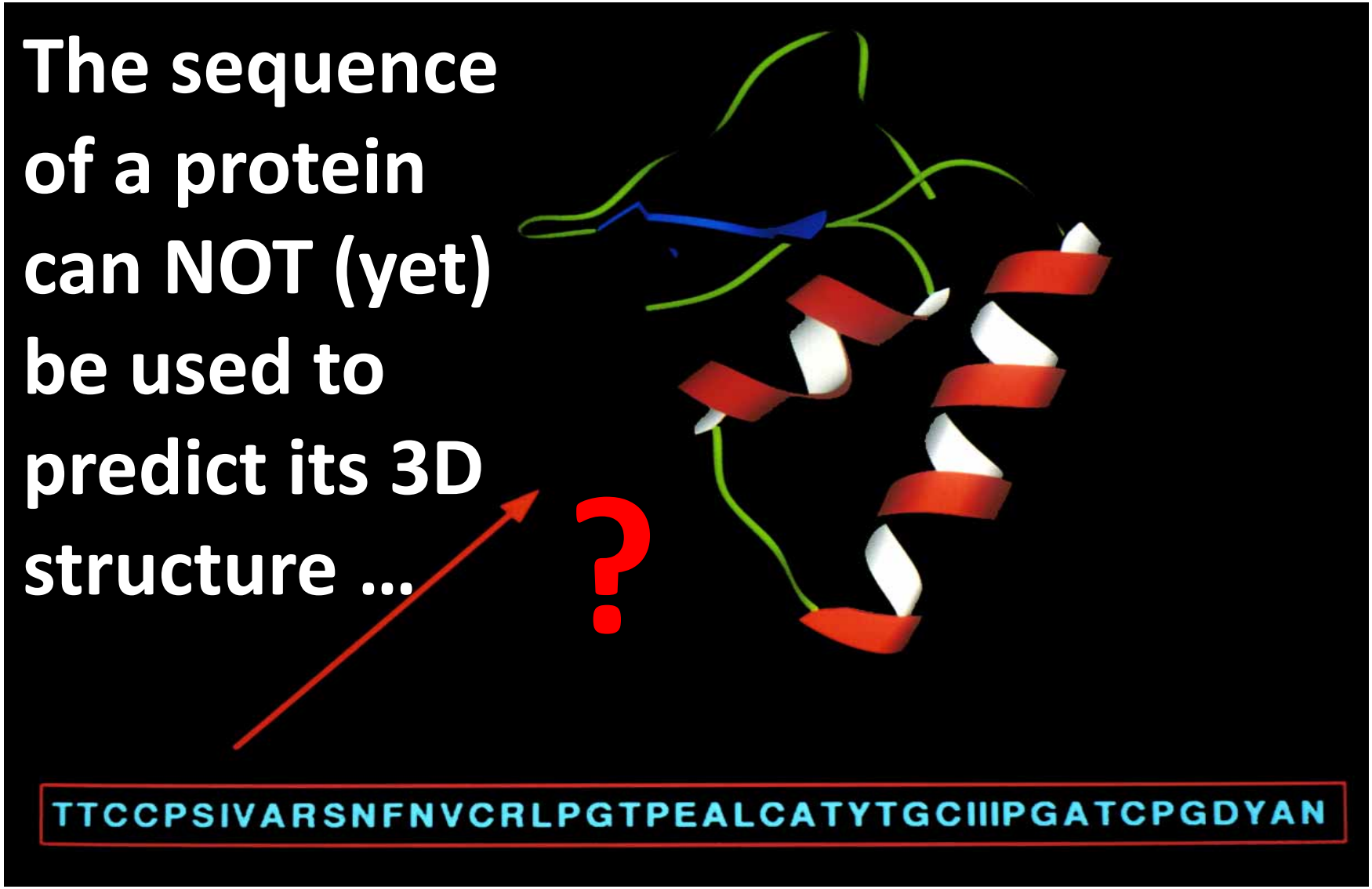
Protein Folding: ὁμολογέω (homologeo)

He, Y., Chen, Y., Alexander, P., Bryan, P. N. & Orban, J. (2008) NMR structures of two designed proteins with high sequence identity but different fold and function. Proceedings of the National Academy of Sciences, 105, 38, 14412.



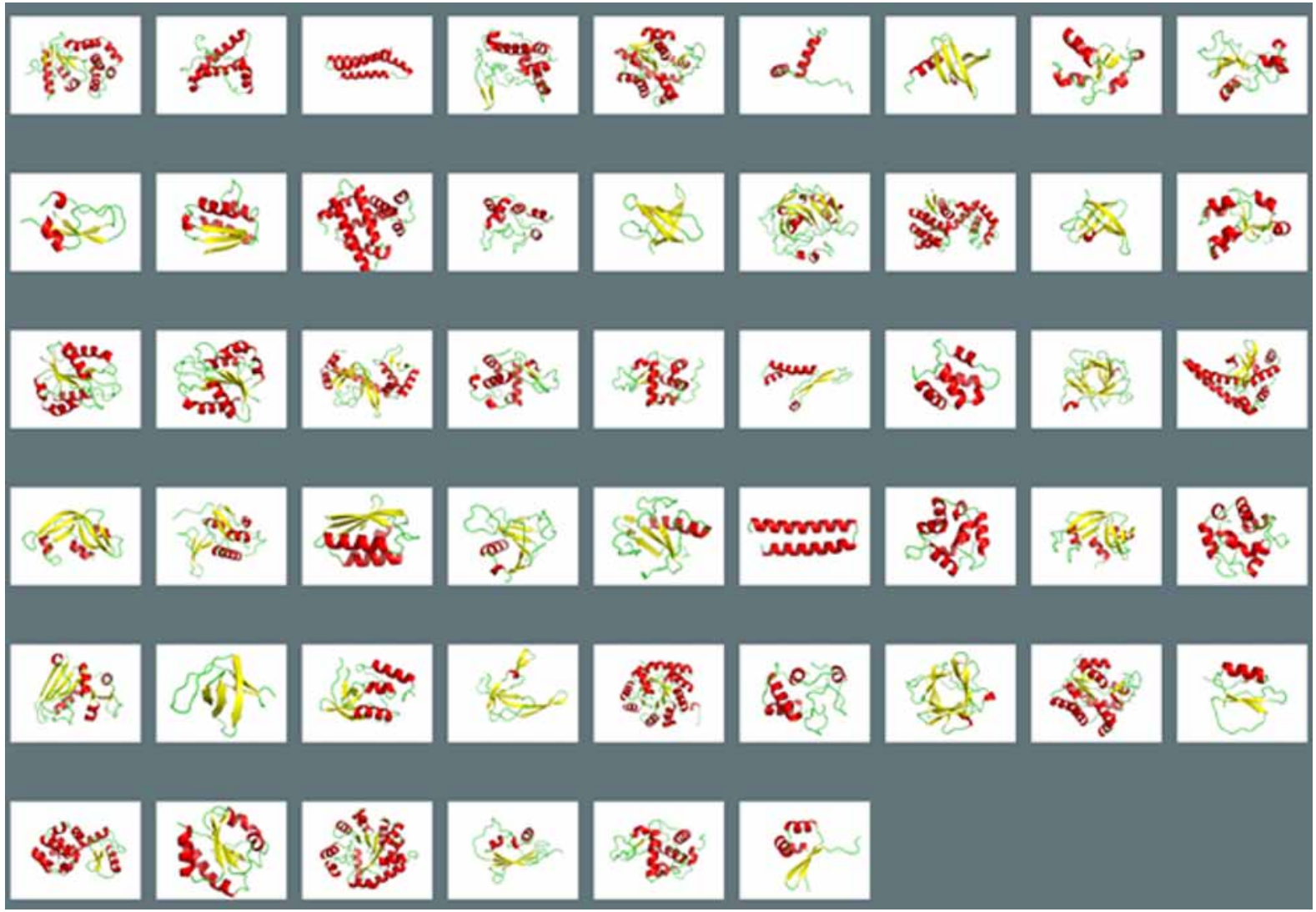

```
T0499   TTYKLILNLKQAKEEAIKEAVDAGTAEKYFKLIANAKTVEGVWTYKDEIKTFTVTE
        ||||||||||||||||||X||||||||X||||||||||||||X||||||||||||
T0498   TTYKLILNLKQAKEEAIKELVDAGTAEKYIKLIANAKTVEGVWTLKDEIKTFTVTE
```

Protein prediction is an old – still unsolved - problem



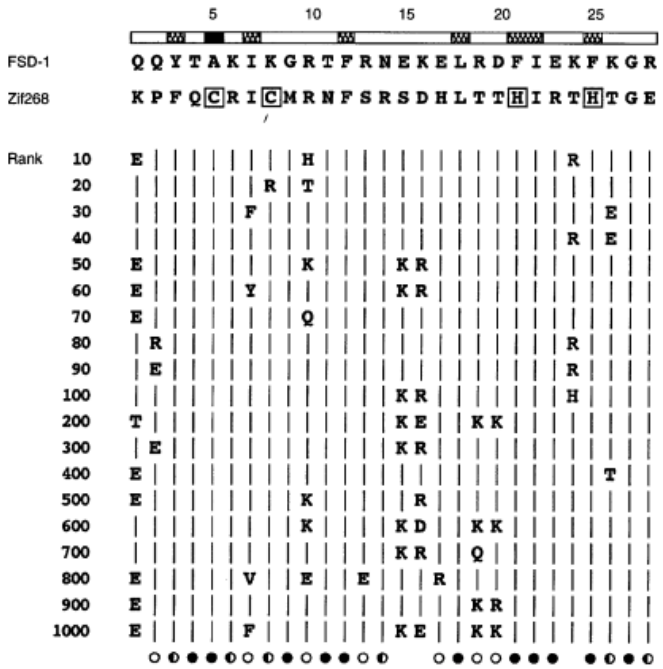
Anfinsen, C. B. **1973.** Principles that Govern the Folding of Protein Chains. Science, 181, (4096), 223-230.

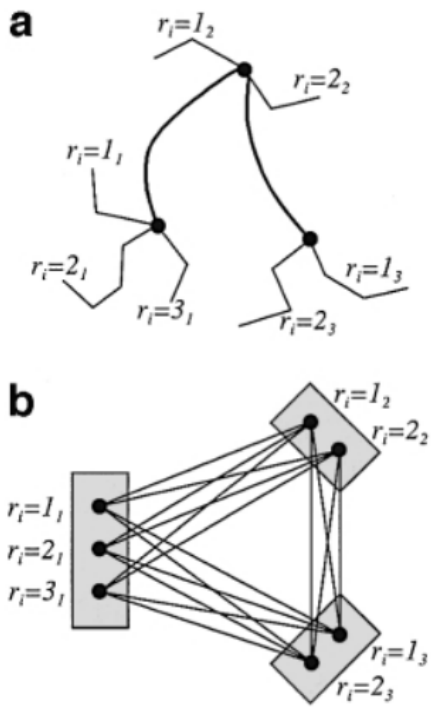
aML still is not able to help sufficiently …



Jia, L., Yarlagadda, R. & Reed, C. C. 2015. Structure Based Thermostability Prediction Models for Protein Single Point Mutations with Machine Learning Tools. Plos One, 10, (9).

Protein Design is a TSP problem

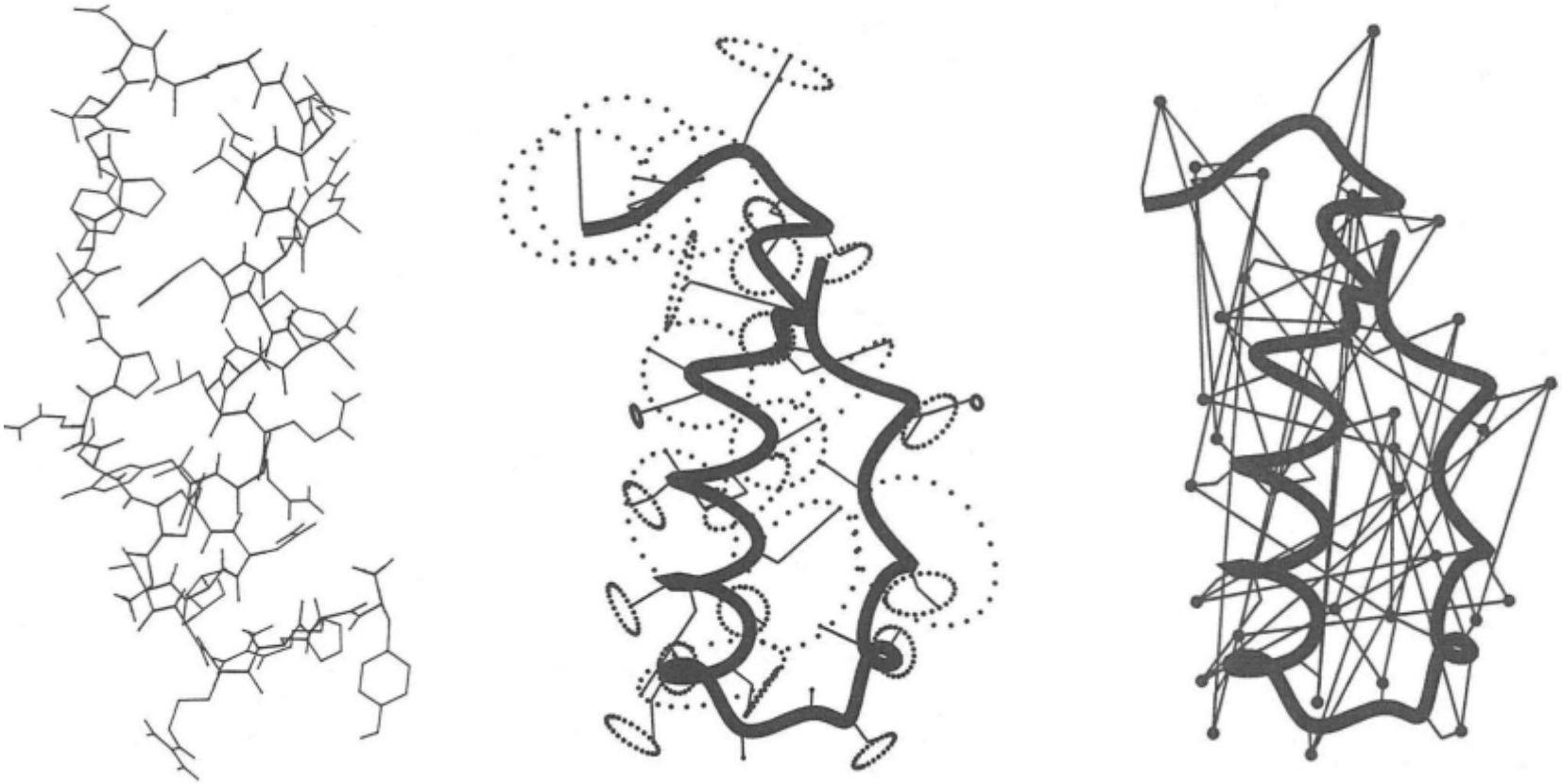




Pierce, N. A. & Winfree, E. 2002. Protein design is NP-hard. Protein Engineering, 15, (10), 779-782.

Dahiyat, B. I. & Mayo, S. L. 1997. De novo protein design: fully automated sequence selection. *Science*, 278, (5335), 82-87.

Protein Folding is a TSP



Bohr, H. & Brunak, S. 1989. A travelling salesman approach to protein conformation. Complex Systems, 3, 9-28

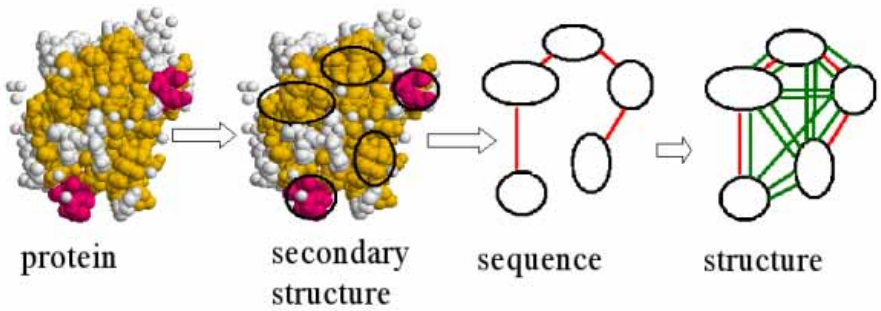
Protein Network Inference

- Hypothesis: most biological functions involve the interactions between many proteins, and the complexity of living systems arises as a result of such interactions.
- In this context, the problem of inferring a global protein network for a given organism,
- **- using all (genomic) data of the organism, is one of the grand challenges in computational biology**

Yamanishi, Y., Vert, J.-P. & Kanehisa, M. 2004. Protein network inference from multiple genomic data: a supervised approach. Bioinformatics, 20, (suppl 1), i363-i370.

Problem: Is Graph Isomorphism NP-complete ?

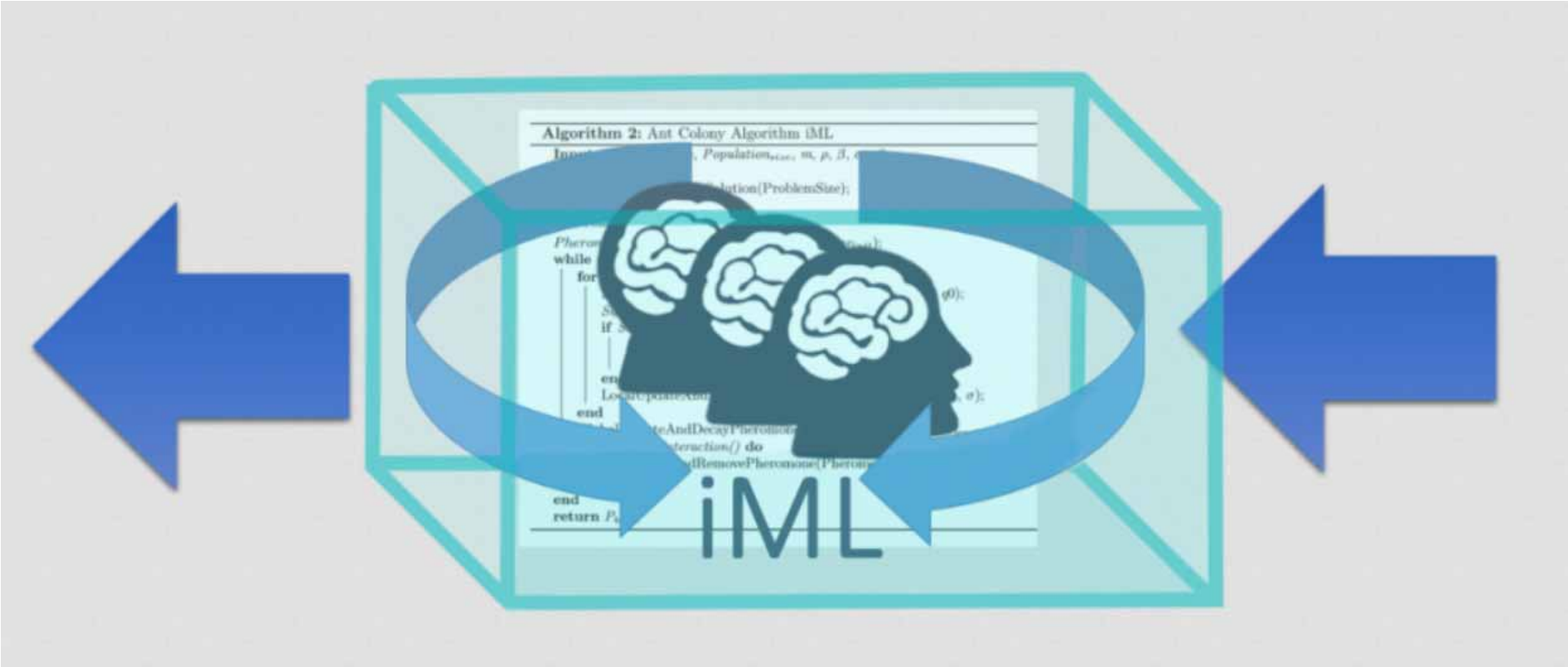
Borgwardt, K. M., Ong, C. S., Schönauer, S., Vishwanathan, S., Smola, A. J. & Kriegel, H.-P. 2005. Protein function prediction via graph kernels. Bioinformatics, 21, (suppl 1), i47-i56.



- Important for health informatics: Discovering relationships between biological components
- Unsolved problem in computer science:
- Can the graph isomorphism problem be solved in polynomial time?
  - So far, no polynomial time algorithm is known.
  - It is also not known if it is NP-complete
  - We know that subgraph-isomorphism is NP-complete

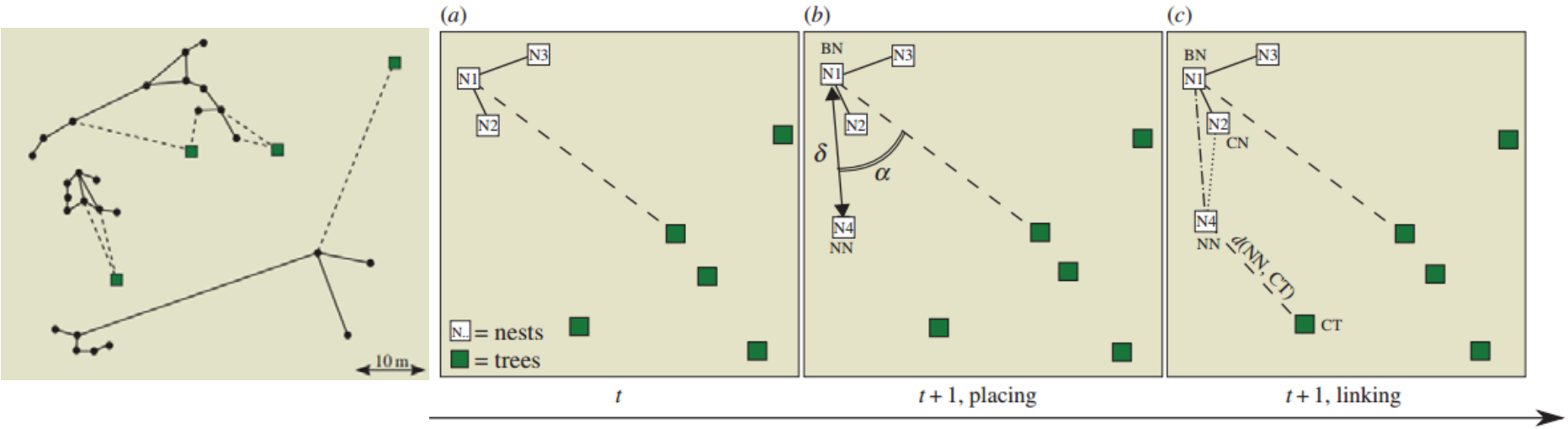
Multi-Agents-in-the-loop

Example: Holzinger Group Project: iML



http://hci-kdd.org/project/iml/

Goal: Finding the shortest path



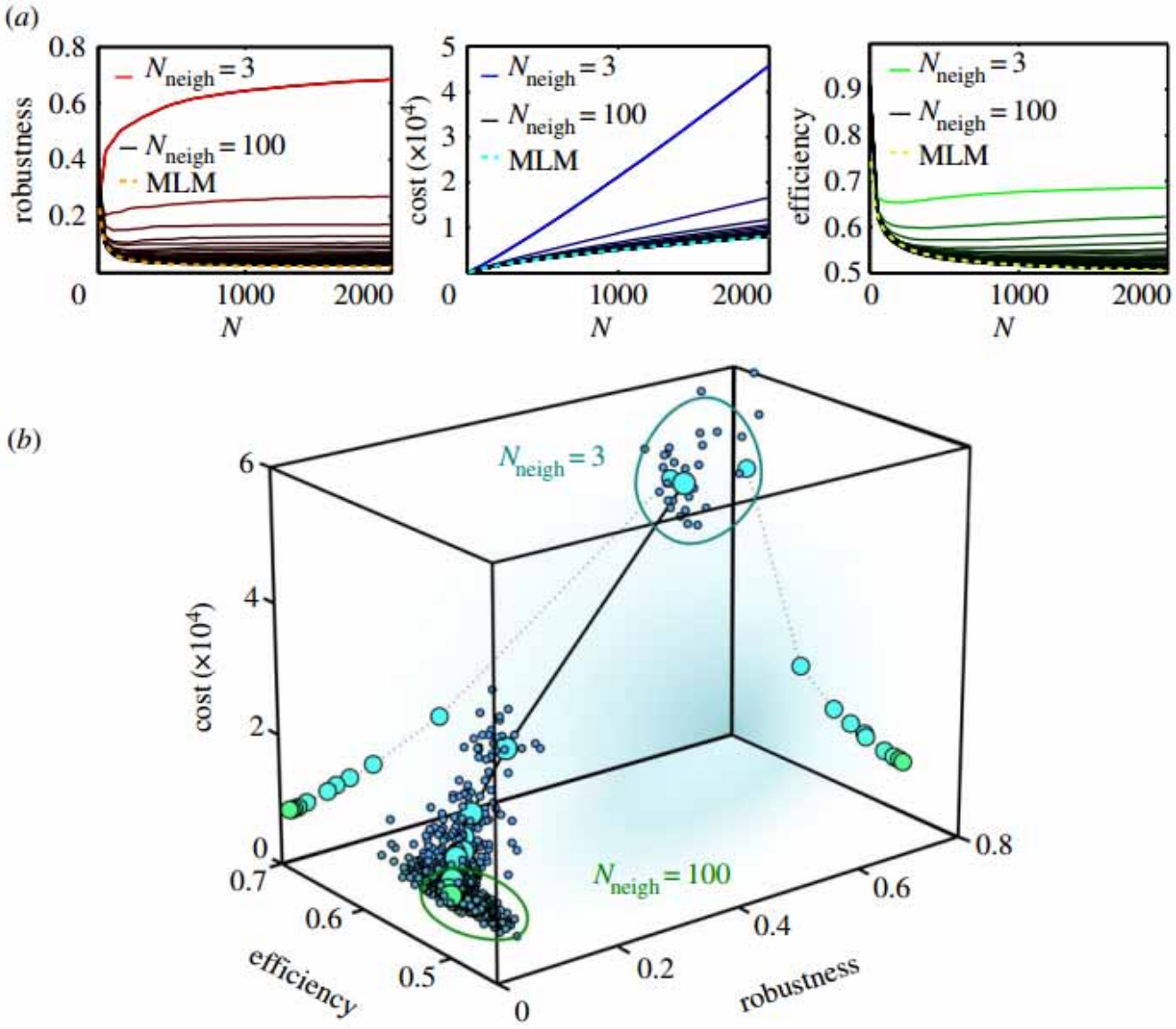
Reasons why ants find the shortest path (minimum linking model):

- 1) Earlier pheromones (the trail is completed earlier)
- 2) More pheromone (higher ant density)
- 3) Younger pheromone (less diffusion)

Soon, the ants will find the shortest path between their home and the food

Bottinelli, A., Van Wilgenburg, E., Sumpter, D. & Latty, T. 2015. Local cost minimization in ant transport networks: from small-scale data to large-scale trade-offs. Journal of The Royal Society Interface, 12, (112), 20150780, doi:10.1098/rsif.2015.0780.

No Free Lunch for the Ants



Bottinelli, A., Van Wilgenburg, E., Sumpter, D. & Latty, T. 2015. Local cost minimization in ant transport networks: from small-scale data to large-scale trade-offs. Journal of The Royal Society Interface, 12, (112), 20150780, doi:10.1098/rsif.2015.0780.

What is the probability for selecting a particular path?

$$p_{ij} = \frac{[\tau_{ij}]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\sum_{l \in J_i^k} [\tau(t)]^{\alpha} \cdot [\eta]^{\beta}}$$

- p_{ij} … **probability** of ants that they, at a particular node i, select the route from node $i \rightarrow j$ (**"heuristic desirability"**)
- $\alpha > 0$ and $\beta > 0$ … the **influence parameters** (α … history coefficient, β …heuristic coefficient) usually $\alpha \approx \beta \approx 2 < 5$
- τ_{ij} … the **pheromone value** for the components, i.e. the amount of pheromone on edge (i, j)
- k … the set of usable components
- J_i … the set of nodes that ant k can reach from v_i (tabu list)
- $\eta_{ij} = \frac{1}{dij}$ … attractiveness computed by a heuristic, indicating the "a-priori **desirability"** of the move

Step 1: Pheromone update

The pheromone on each edge is updated as:

$$\tau_{ij} = (1-\rho)\cdot\tau_{ij} + \Delta\tau_{ij}$$

With:

- ρ : the evaporation rate of the 'old' pheromone
- $\Delta\tau_{ij}$: the 'new' pheromone that is deposited by all ants on edge (i,j) calculated as:

$$\Delta\tau_{ij} = \sum_{k=0}^{m} \Delta\tau_{ij}^{k}$$

Step 2: Pheromone update

The pheromone that is deposited on edge (i,j) by ant k is calculated as:

$$\Delta\tau_{ij}^{k} = \begin{cases} Q / L_k & if\ (i,j) \in T_k \\ 0 & otherwise \end{cases}$$

With:

- Q : a heuristic parameter
- T_k : the path traversed by ant k
- L_k : the length of T_k calculated as the sum of the lengths of all the edges of T_k

Now with the human in the loop

```
Input  : ProblemSize, m, β, ρ, σ, q0
Output: Pbest
Pbest ← CreateHeuristicSolution(ProblemSize);
Pbest_cost ← Cost(Pbest);
Pheromone_init ← 1.0 / (ProblemSize × Pbest_cost);
Pheromone ← InitializePheromone(Pheromone_init);
while ¬StopCondition() do
    for i = 1 to m do
        S_i ← ConstructSolution(Pheromone, ProblemSize, β, q0);
        Si_cost ← Cost(S_i);
        if Si_cost ≤ Pbest_cost then
            Pbest_cost ← Si_cost;
            Pbest ← S_i;
        end
        LocalUpdateAndDecayPheromone(Pheromone, S_i, Si_cost, ρ);
    end
    GlobalUpdateAndDecayPheromone(Pheromone, Pbest, Pbest_cost, ρ);
    while isUserInteraction() do
        GlobalAddAndRemovePheromone(Pheromone, Pbest, Pbest_cost, ρ);
    end
end
return P_best;
```

Holzinger, A., Plass, M., Holzinger, K., Crisan, G., Pintea, C. & Palade, V. 2016. Towards interactive Machine Learning (iML): Applying Ant Colony Algorithms to solve the Traveling Salesman Problem with the Human-in-the-Loop approach. Springer Lecture Notes in Computer Science LNCS 9817. 81-95, doi:10.1007/978-3-319-45507-56.

Use of heuristic information

- The attractiveness η_{ij} of edge (i, j) is computed by a heuristic, indicating the a-priori desirability of that particular move
- The pheromone trail level τ_{ij} of edge (i, j) indicates how proficient it was in the past
- $\alpha = 0$ is a greedy approach and $\beta = 0$ represents the selection of tours that may not be optimal
- Consequently, we speak of a "trade-off" between speed and quality

Results from a first (small-scale) experiment

Table 1

	Mean		Standard Deviation		Shortest Path	
	Computer	iML	Computer	iML	Computer	iML
Problem 1	4499,34	4614,01	109,61	232,35	4301,20	4271,48
Problem 2	3496,92	3679,37	40,15	191,75	3407,50	3450,56
Problem3	4362,35	4563,33	25,34	293,84	4320,06	4250,90

Table 1: In mean the computer has the smaller mean and standard deviation. By finding the shortest path the iML approach gives better results (Additional Boxplots in the other sheets)

Table 2

	Number of Stops	Number of Interactions	Number of Datasets
Problem 1	2,58	8,7	24
Problem 2	3	16,15	13
Problem3	2,5	22,17	12

Table 2: Overview User interactions

- Evidence for the human-in-the-loop, however, strongly dependent on end-user expertise;
- Further experiment will be camouflaged as a game, to provide the same start criteria for everybody

Holzinger, A., Plass, M., Holzinger, K., Crisan, G., Pintea, C. & Palade, V. 2016. Towards interactive Machine Learning (iML): Applying Ant Colony Algorithms to solve the Traveling Salesman Problem with the Human-in-the-Loop approach. Springer Lecture Notes in Computer Science LNCS 9817. Heidelberg, Berlin, New York: Springer, pp. 81-95

Online iML-Experiment still running – large scale

- You can have a look at the online Experiment:
- http://hci-kdd.org/projects/iml-experiment

Experiment: Interactive Machine Learning for the Traveling-Salesman-Problem

07 Conclusion

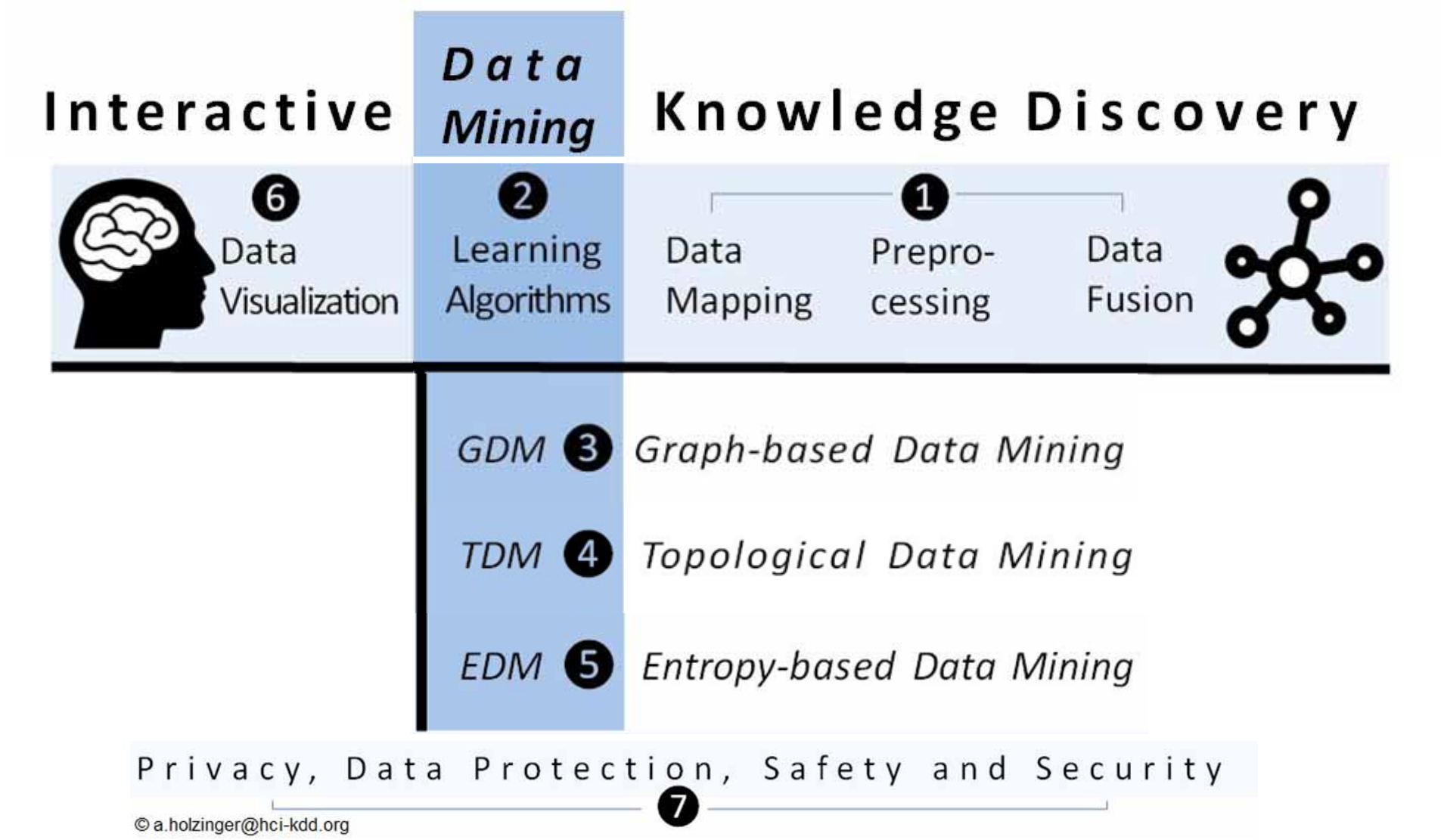
Successful Machine Learning needs a …

concerted effort

international

without boundaries …

The HCI-KDD machine learning pipeline



Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning. IEEE Intelligent Informatics Bulletin, 15, (1), 6-14.

- Application of aML in complex domains seems elusive in the near future, a good example are Gaussian processes, where aML (e.g., kernel machines) struggle on function extrapolation - which is trivial for human learners.
- Making use of human cognitive abilities can be used to solve problems, when lacking big data sets, on complex data and/or rare events, where traditional learning algorithms suffer due to insufficient training samples.
- Here human expertise can assist in solving problems which otherwise would remain NP-hard.
- Successful application of ML requires a concerted effort of (1) data science, (2) algorithms, (3) graphs, (4) topology, (5) entropy, (6) data visualization, and (7) privacy, data protection, safety, and security—following the HCI-KDD approach in combining the best of two worlds.

08 Appendix

Reading (Nachlese)

- Holzinger, A. 2016. Interactive Machine Learning for Health Informatics: When do we need the human-in-the-loop? Springer Brain Informatics (BRIN), 3, (2), 119-131, doi:10.1007/s40708-016-0042-6.
- Holzinger, A. 2016. Interactive Machine Learning (iML). Informatik Spektrum, 39, (1), 64-68, doi:10.1007/s00287-015-0941-6.
- Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning. IEEE Intelligent Informatics Bulletin, 15, (1), 6-14.
- Holzinger, A. 2014. Extravaganza Tutorial on Hot Ideas for Interactive Knowledge Discovery and Data Mining in Biomedical Informatics. In: Slezak, D., Tan, A.-H., Peters, J. F. & Schwabe, L. (eds.) Brain Informatics and Health, BIH 2014, Lecture Notes in Artificial Intelligence, LNAI 8609. Heidelberg, Berlin: Springer, pp. 502-515, doi:10.1007/978-3-319-09891-3_46.
- Holzinger, A. 2013. Human–Computer Interaction and Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together? In: Cuzzocrea, A., Kittl, C., Simos, D. E., Weippl, E. & Xu, L. (eds.) Multidisciplinary Research and Practice for Information Systems, Springer Lecture Notes in Computer Science LNCS 8127. Heidelberg, Berlin, New York: Springer, pp. 319-328, doi:10.1007/978-3-642-40511-2_22.