



**PennState**  
College of Engineering



**CHOT**

# *Machine learning applications: Towards personalized medicine and biomechanical modeling.*

**Christian E. López<sup>1</sup> & Dr. Conrad S. Tucker<sup>1,2</sup>**

<sup>1</sup> Department of Industrial and Manufacturing Engineering,

<sup>2</sup> School of Engineering Design Technology and Professional Programs  
The Pennsylvania State University, University Park





*The title of the first paper most frequently cited is:  
"Some Studies in Machine Learning Using the Game of Checkers"*

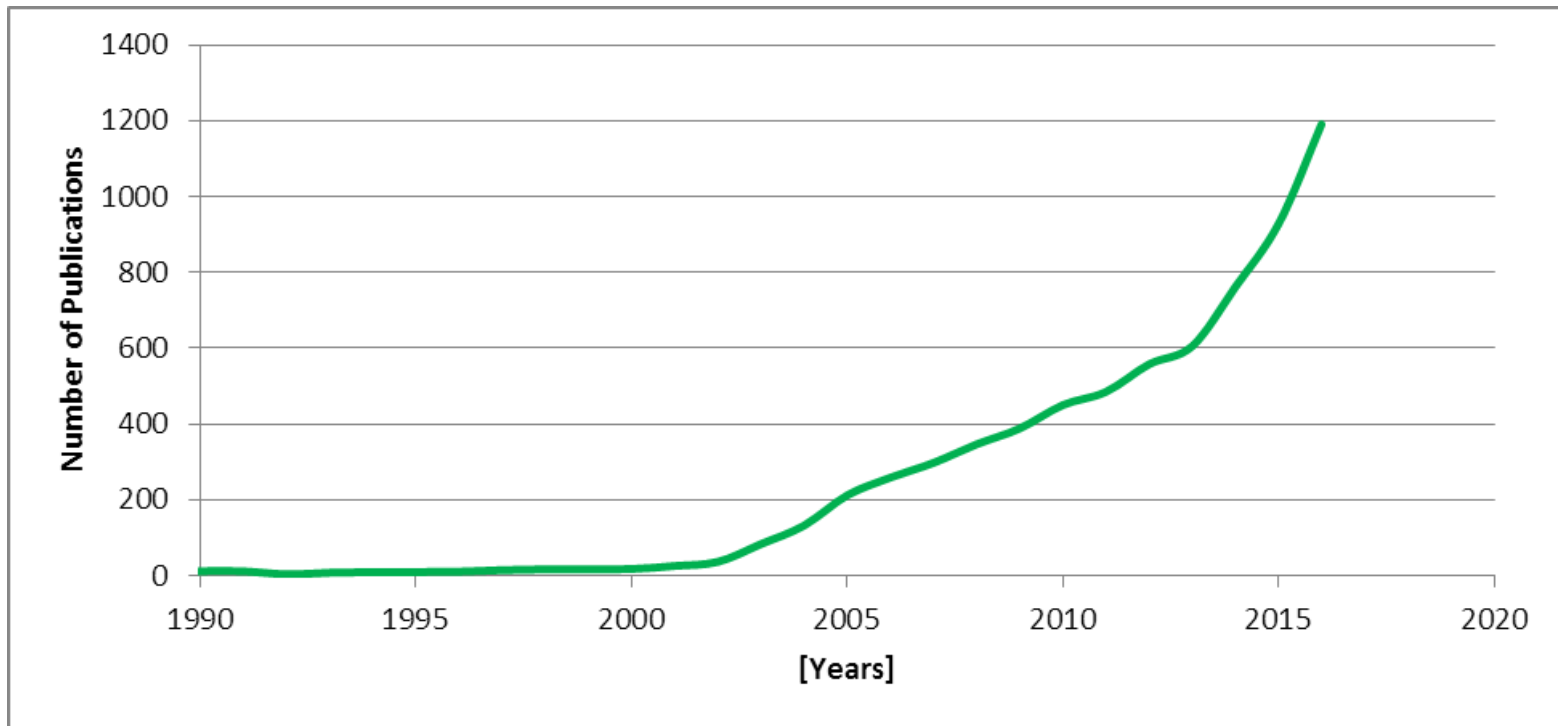


# Biomedical and Bioengineering fields are gaining interest in Machine Learning methods



*“Biomedical Machine Learning OR Bioengineering Machine Learning”*

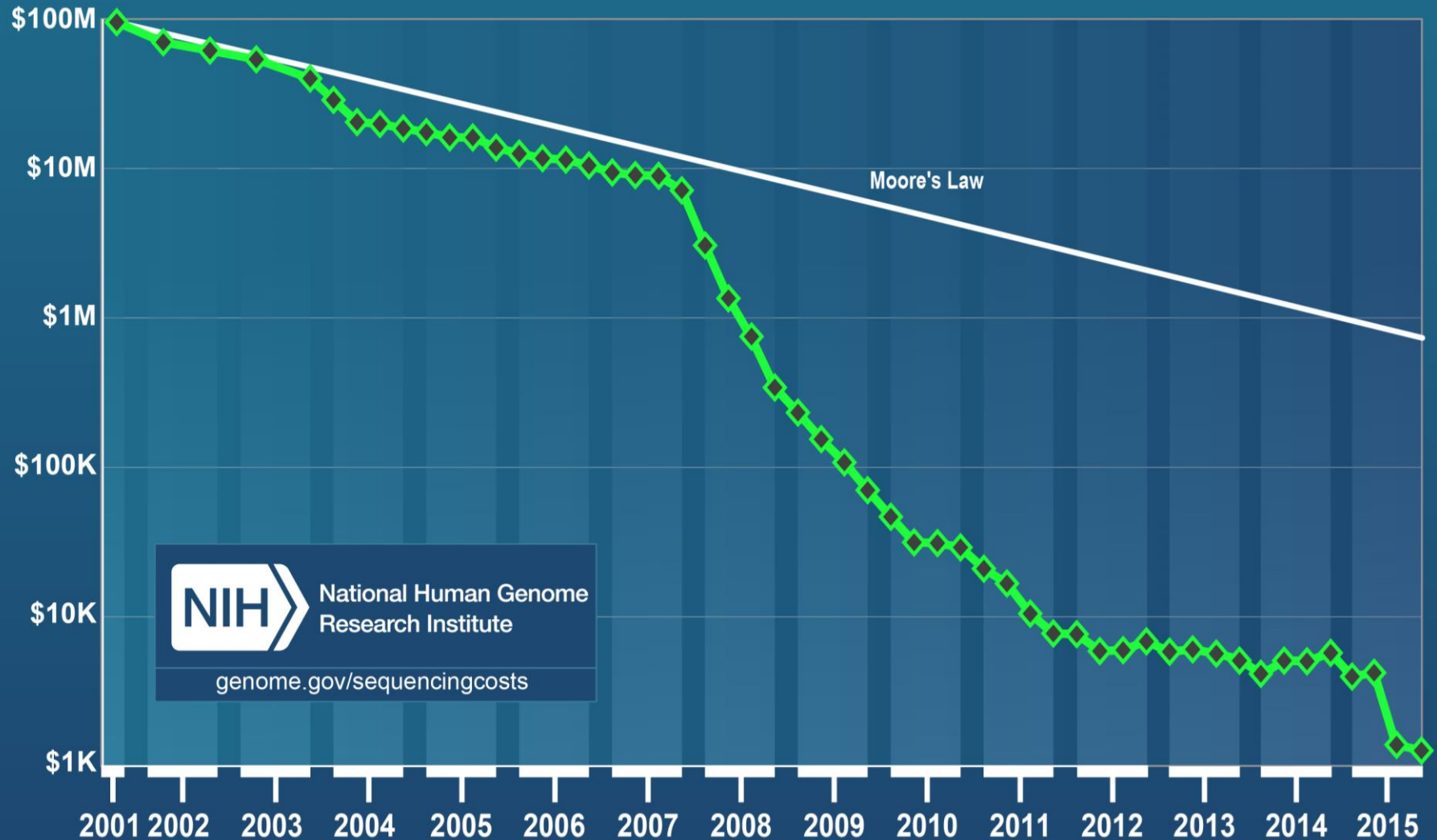
**7976 records**



*Extracted 10/29/2017*

# The cost of generating a human genome sequence is continuously decreasing, as technology advances

[NIH, 2016]

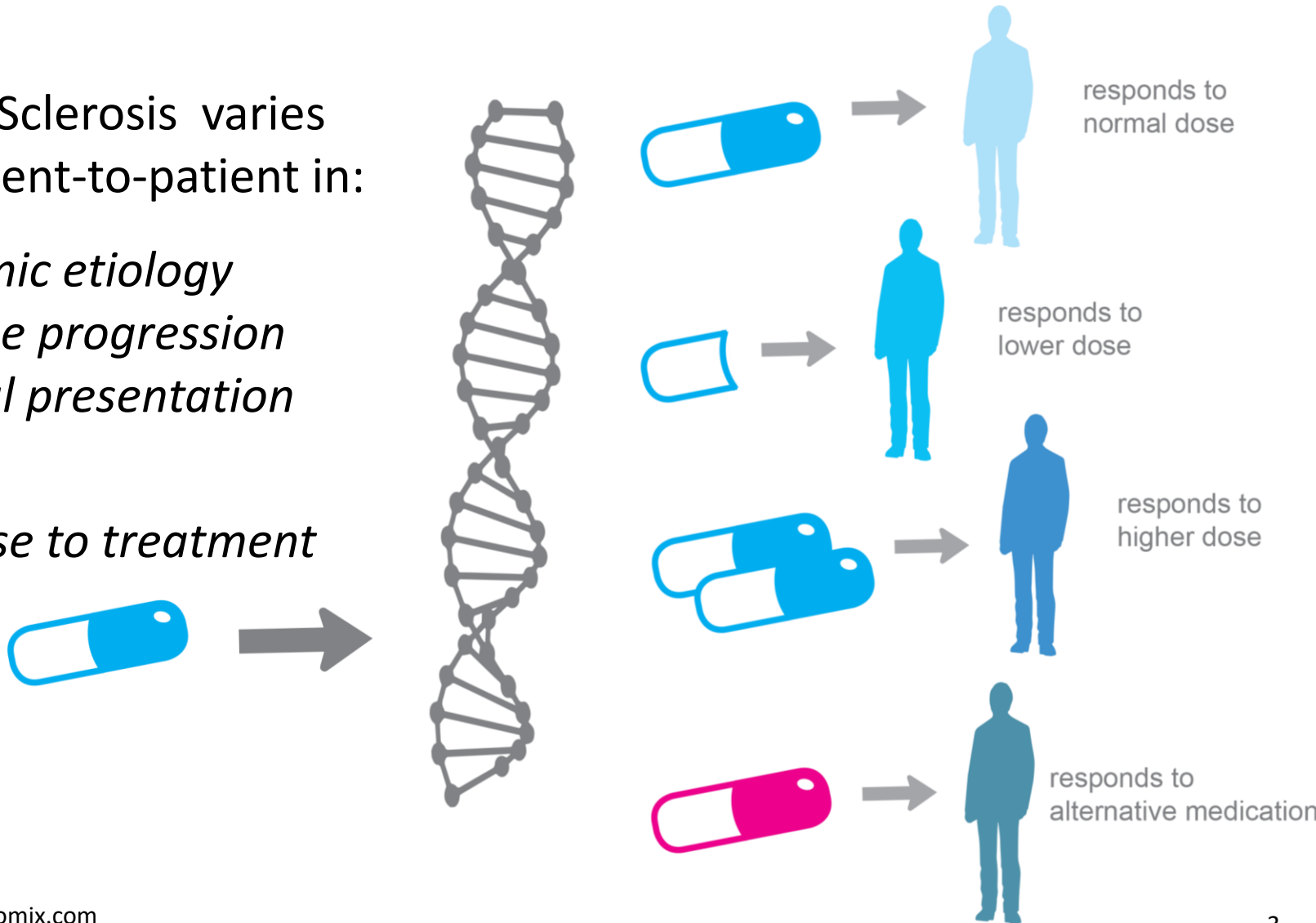


# Personalized medicine to improve clinical outcomes by incorporating patient-specific data

[Kulakova et al., 2014]

Multiple Sclerosis varies from patient-to-patient in:

- *Genomic etiology*
- *Disease progression*
- *Clinical presentation*
- *Response to treatment*

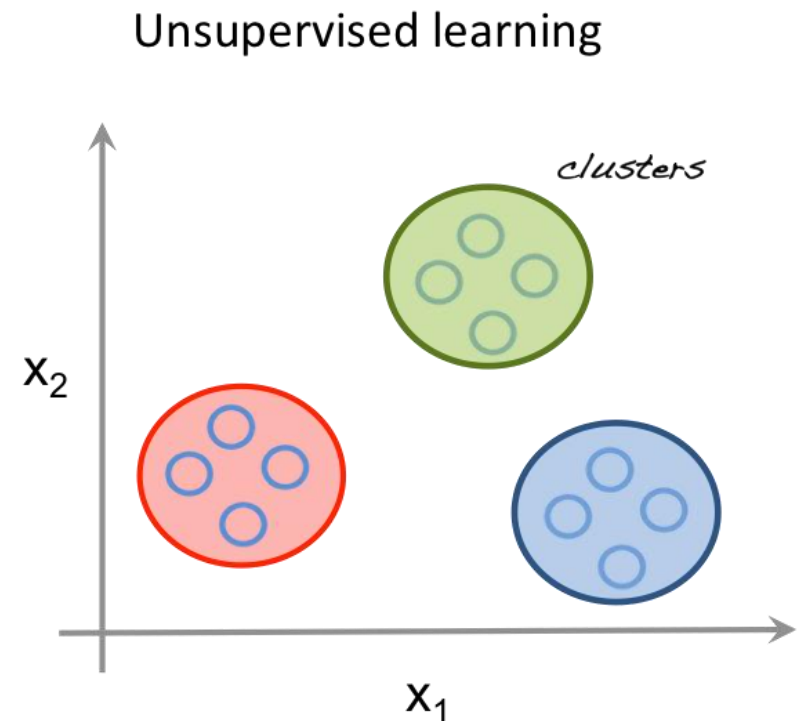


# What if there are not class labels available, the labels are incorrect, or the labels are not specific enough?

*“If we are interested in discovering what types of labels best explain the data ...we must use unsupervised rather than supervised learning”*

[Libbrecht & Noble, 2015, pag 4]

- ✓ Identify potentially interesting new cluster structures
- ✓ Implement without class label data

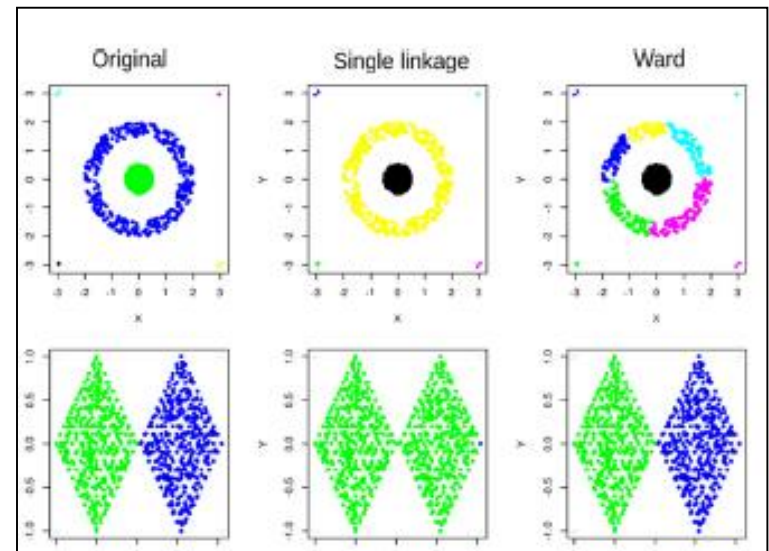
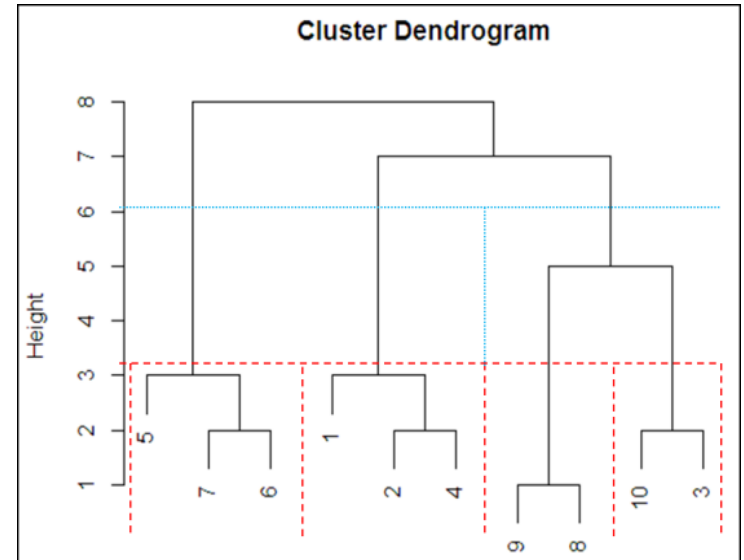


# Hierarchical clustering has shown to consistently outperform other algorithms, in biomedical datasets

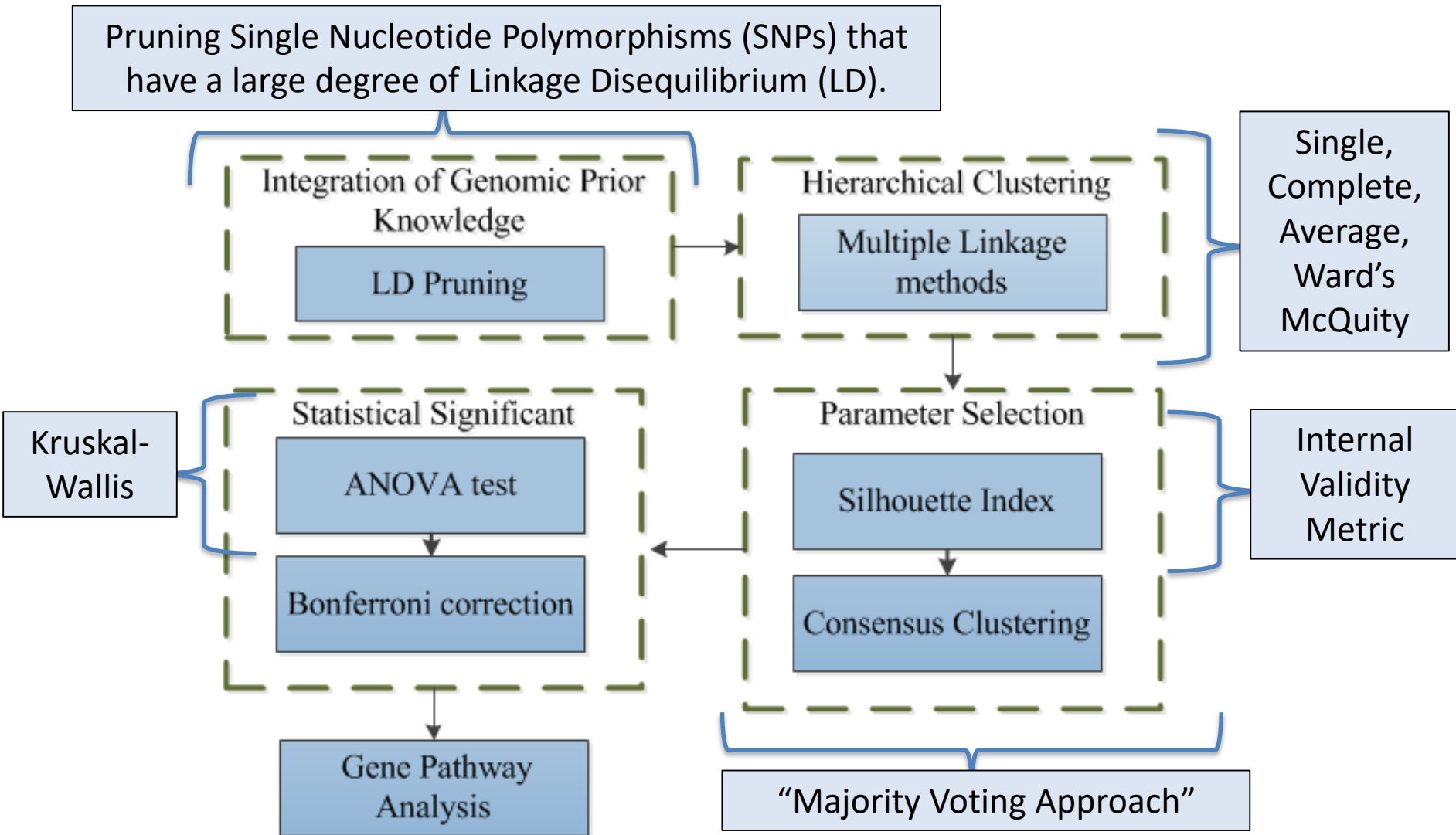
[Wiwie et al., 2015]

- ✓ Discover several layers of clustering structure
- ✓ Easy visualization

- ❑ Predetermine parameters (e.g., number of cluster)
- ❑ Test for Cluster Significance
- ❑ Different Methods provide different results



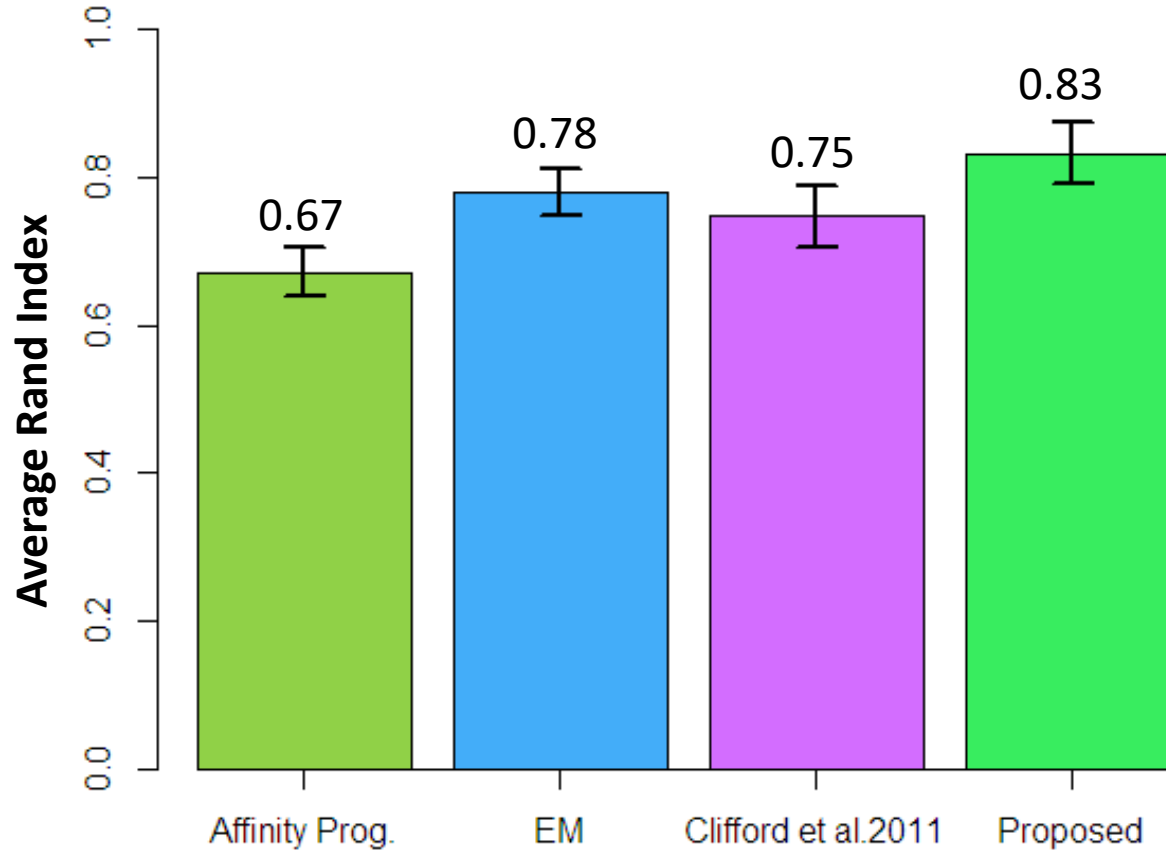
# Unsupervised machine learning method to cluster patients based on their genomic makeup





# Our method outperforms other methods, based on initial Benchmark results

*“Measurement of agreement between Cluster sets”*



**Perfect Clustering:**

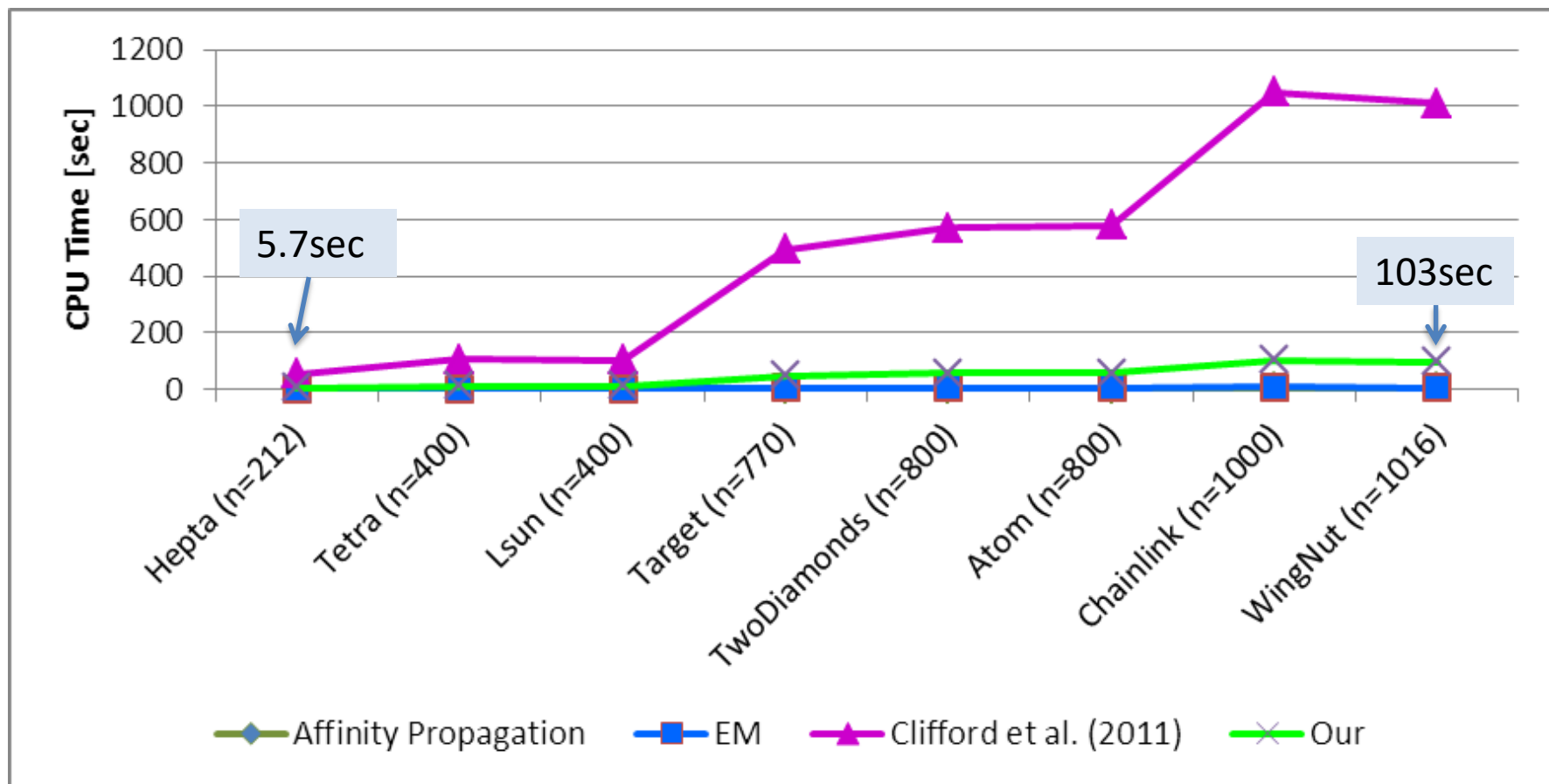
12.5%

25%

25%

50%

# Our method provides the best results in real-time (e.g., <2min)



Average time:

2.2sec

3.7sec

495.9sec

48.2sec

The method can be used to identify genetically unique subpopulations within an patient cohort

***Implications for:***

- ❖ Early disease detection in at-risk populations
- ❖ Predicting disease course
- ❖ Predicting response to treatment
- ❖ Identifying specific groups to target for future study

The method can be used to identify genetically unique subpopulations within an patient cohort

**Future work will focus on improving the benchmark analysis, and testing in other GWAS datasets**



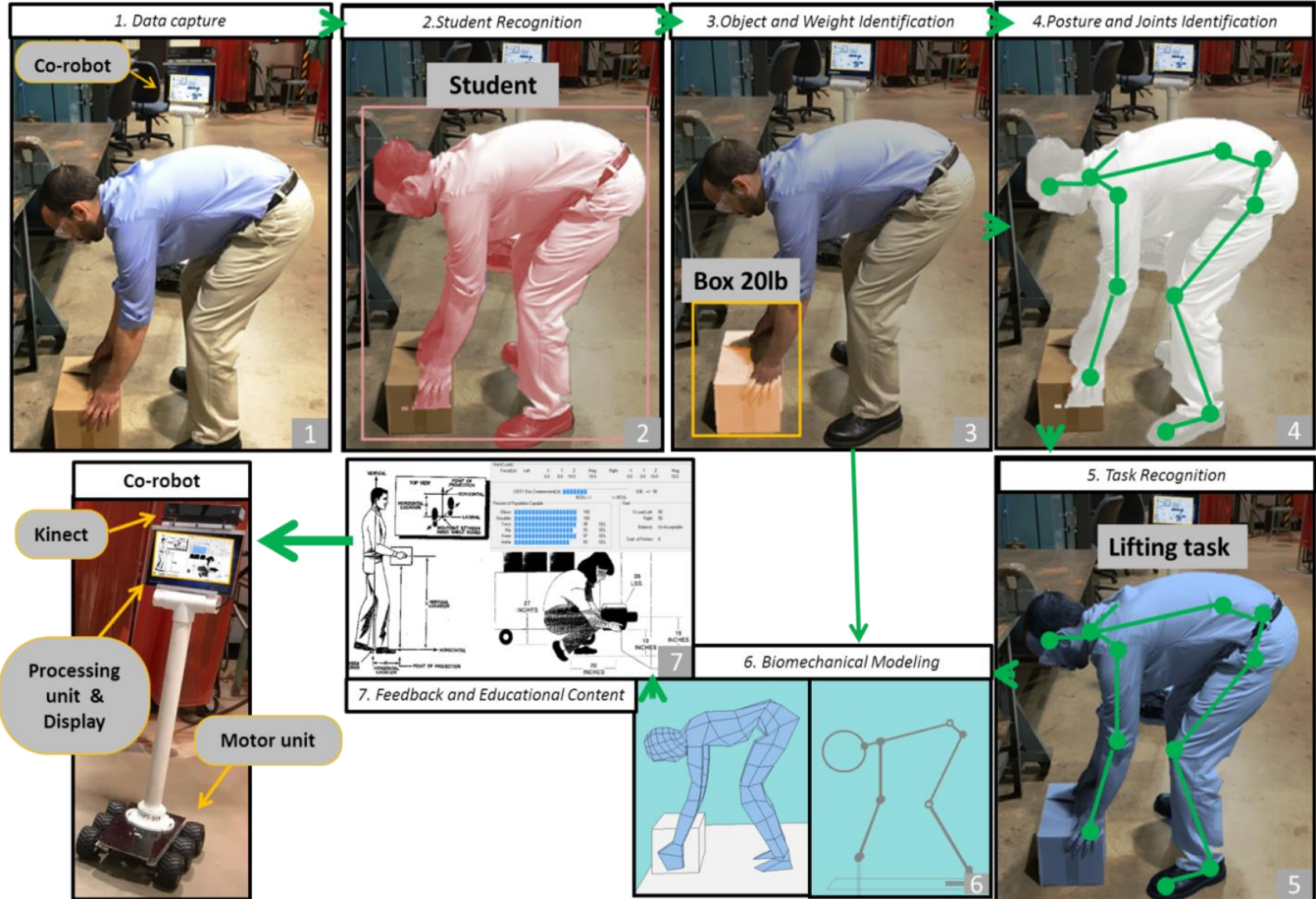
# Advanced Personalized Medicine



# Advanced Personalized Biomechanical Modeling

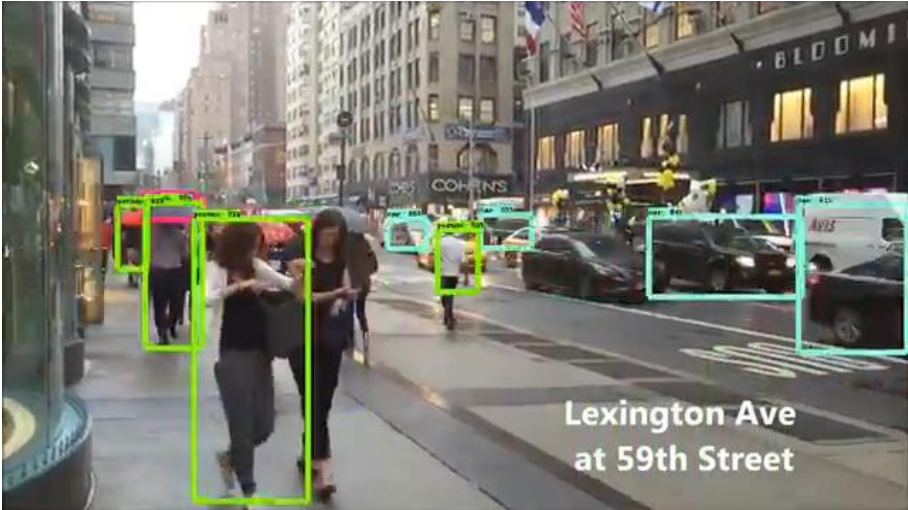
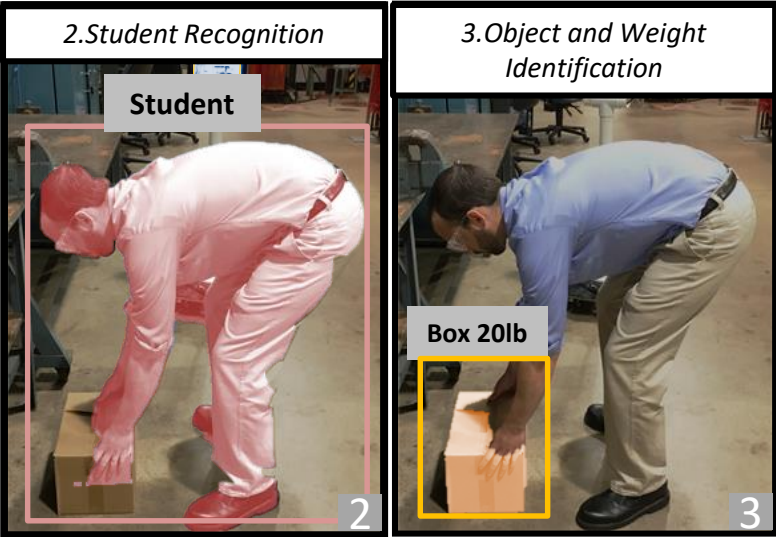
# Machine Learning and computer vision technology can be used to provide real-time ergonomic feedback

[Lopez & Tucker , 2017]



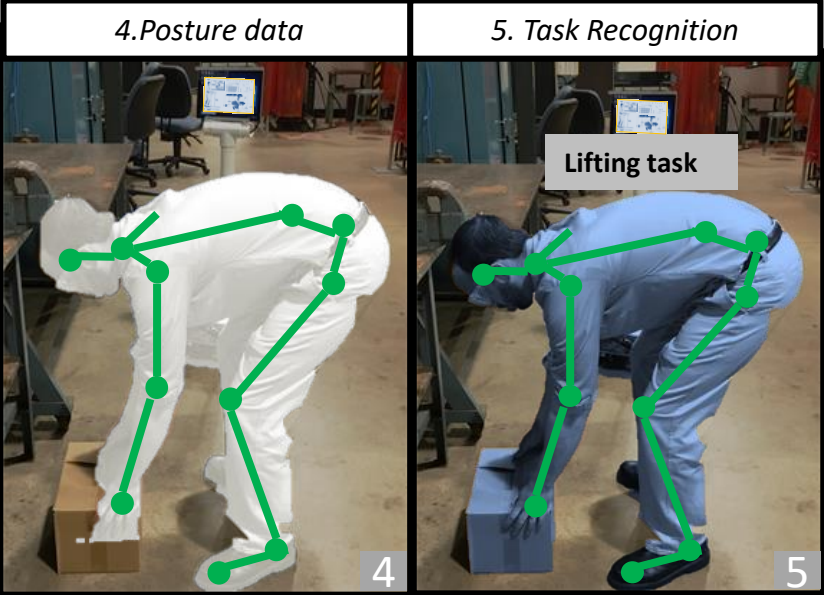
# There has been significant advancements in Computer Vision and Object Recognition

[Dering and Tucker, 2017]



[Google, 2017]

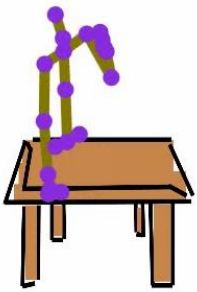
# There has been significant advancements in Machine Learning Applications



[Cao et al., 2016]

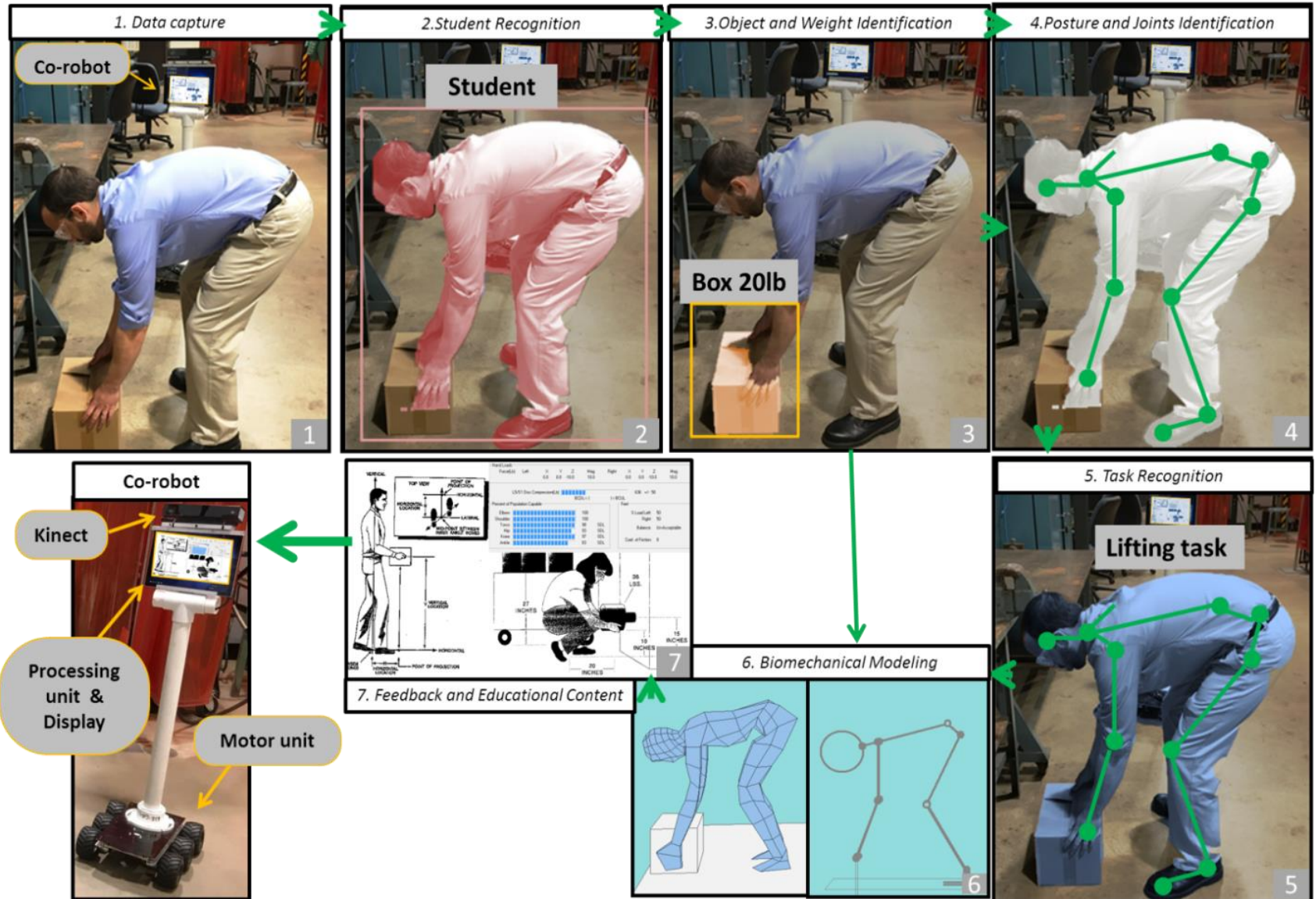


[Dering and Tucker, 2017]

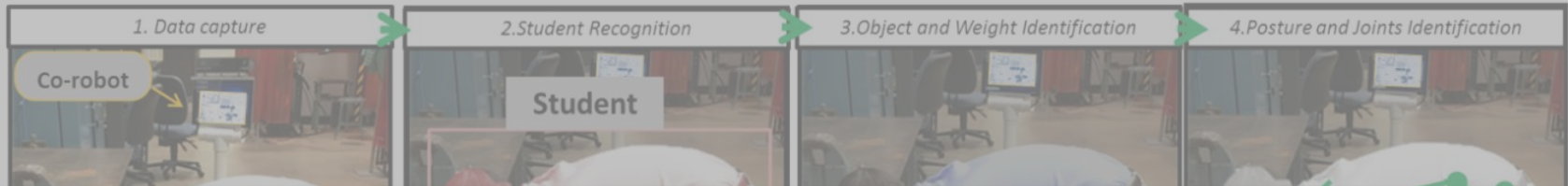




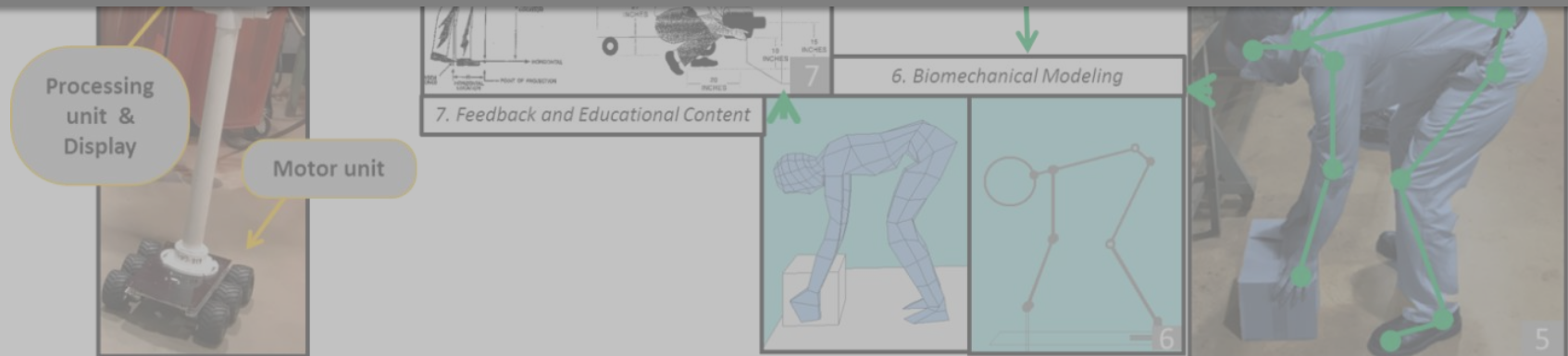
# These technologies could be integrated to provide real-time ergonomic feedback.



These technologies could be integrated to provide real-time ergonomic feedback.



Even though this method is in its initial stages of design and needs to be further tested, its potential capabilities are promising.





**PennState**  
College of Engineering



**D.A.T.A.**  
**Design Analysis**  
Technology Advancement Laboratory

# Thank you!

**Scott Tucker & Tarik Salameh**

MD/PhD Student Hershey College of Medicine



**CHOT**

[www.chot.psu.edu](http://www.chot.psu.edu)

*This research is funded in part by NSF NRI # 1527148, and NSF I/UCRC Center for Healthcare Organization Transformation (CHOT), NSF I/UCRC grant #1067885. Any opinions, findings, or conclusions found in this paper are those of the authors and do not necessarily reflect the National Science Foundation.*

Samuel, Arthur L. (1959). "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development*. 3(3): 210-229

NIH-National Human Genome Research Institutes (2016). The Cost of Sequencing a Human Genome. <https://www.genome.gov/sequencingcosts/> (10/31/2017)

O. Kulakova, E. Tsareva, D. Lvovs, A. Favorov, A. Boyko, and O. Favorova, "Comparative pharmacogenetics of multiple sclerosis: INF-B versus glatiramer acetate," *Pharmacogenomics*, vol. 15, no. 5, pp. 679–85, 2014.

M. W. Libbrecht and W. S. Noble, "Machine learning in genetics and genomics," *Nature Reviews. Genetics*, vol. 16, no. 6, pp. 321–332, 2015.

M. W. Libbrecht and W. S. Noble, "Machine learning in genetics and genomics," *Nature Reviews. Genetics*, vol. 16, no. 6, pp. 321–332, 2015.

C. Wiwie, J. Baumbach, and R. Röttger, "Comparing the performance of biomedical clustering methods," *Nature Methods*, vol. 12, no. 11, pp. 1033–1038, 2015.

H. Clifford, F. Wessely, S. Pendurthi, and R. D. Emes, "Comparison of clustering methods for investigation of genome-wide methylation array data," *Frontiers in Genetics*, vol. 2, no. DEC, pp. 1–11, 2011.

Lopez, C. E., & Tucker, C. S. (2017). Towards Real-time Ergonomics Feedback and Educational Content with the use of Co-Robots. In the 2017 ASEE Mid-Atlantic Conference

Dering M, Tucker C, Kumara S. An Unsupervised Machine Learning Approach To Assessing Designer Performance During Physical Prototyping. *ASME J Comput Inf Sci Eng*. 2017. doi:10.1115/1.4037434.

Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2016). Realtime multi-person 2d pose estimation using part affinity fields. *arXiv preprint arXiv:1611.08050*.