Application of Artificial Intelligence in Health Informatics

Prof. Gajendra P.S. Raghava

Head, Department of Computational Biology

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Web Site:
http://webs.iiitd.edu.in/raghava/

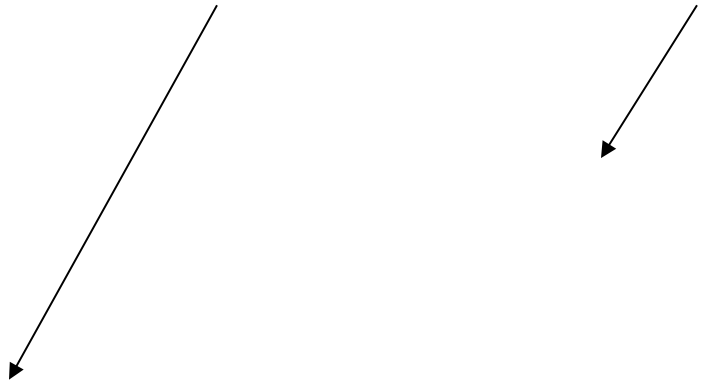
Managing Big Data

Classification of Digital Data

- Structured data (10%)
  - RDBMS (MySql, PostgreSQL, Oracle, DB2 etc.)
  - Spreadsheet
  - On-line transaction process
- Semi-structured data (10%)
  - XML (eXtensible Markup Language); JSON (Java Script Object Notation); other markup language.
- Unstructured data (80%)
  - Data mining; Natural language processing (NLP); Text analytics; Noisy text analytics

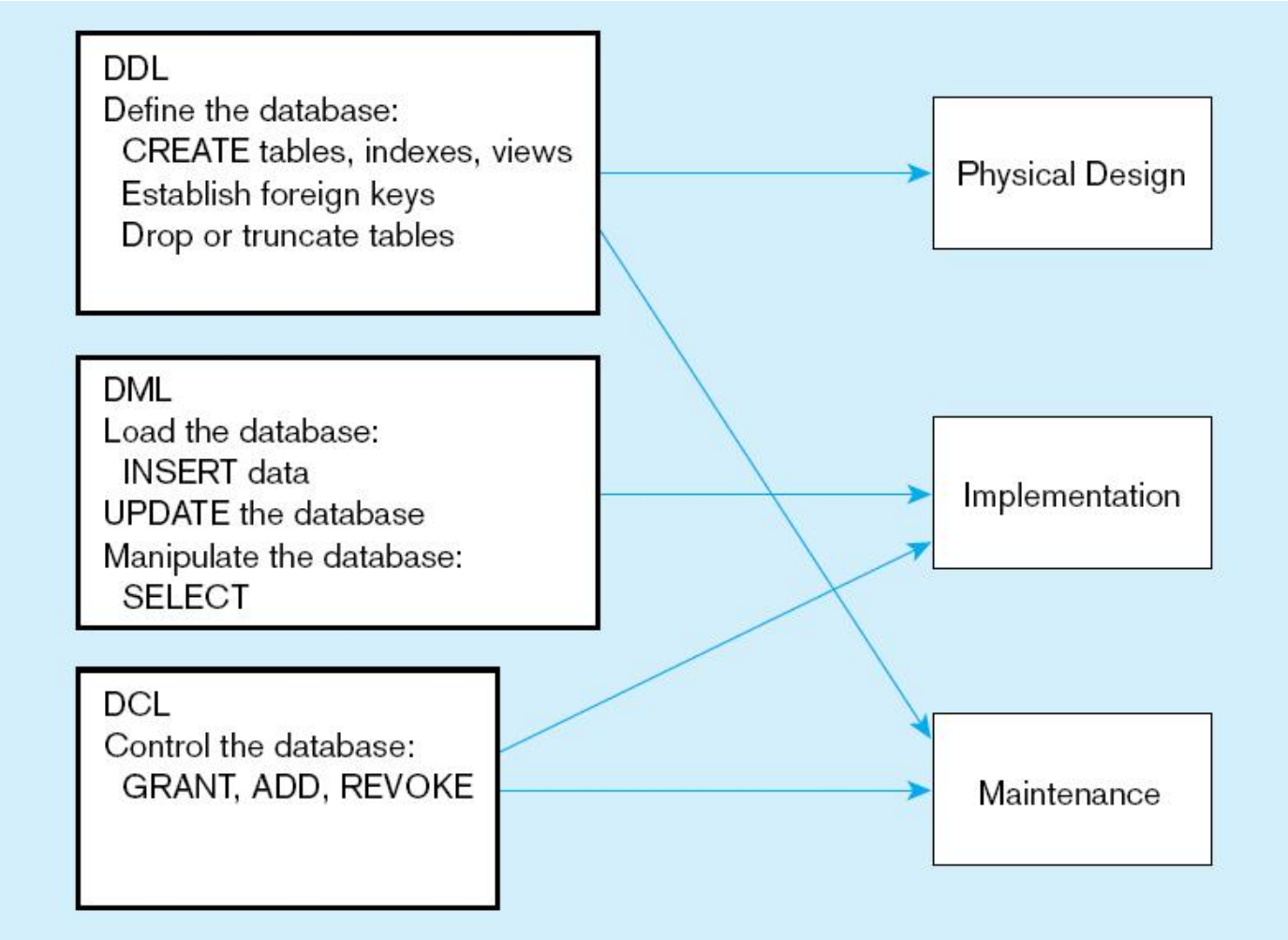
Relational Database Management System (RDBMS)

- Most successful model
  - Tables for representing data
  - Flexible logical structure for data representation
  - A series of row/column
- Disadvantages
  - Too much overhead to maintain uniformity
  - Ease-of-use allows careless use of RDBMS



Structure Query Language

- Data Definition Language (DDL)
  - Create/alter/delete tables and their attributes
- Data Manipulation Language (DML)
  - Query one or more tables – discussed
  - Insert/delete/modify tuples in tables
- Data Control Language (DCL)
- SQL data is case-sensitive,
- SQL commands are not
- Schema
  - The structure that contains descriptions of objects created by a user (base tables, views, constraints)
- Catalog: A set of schemas



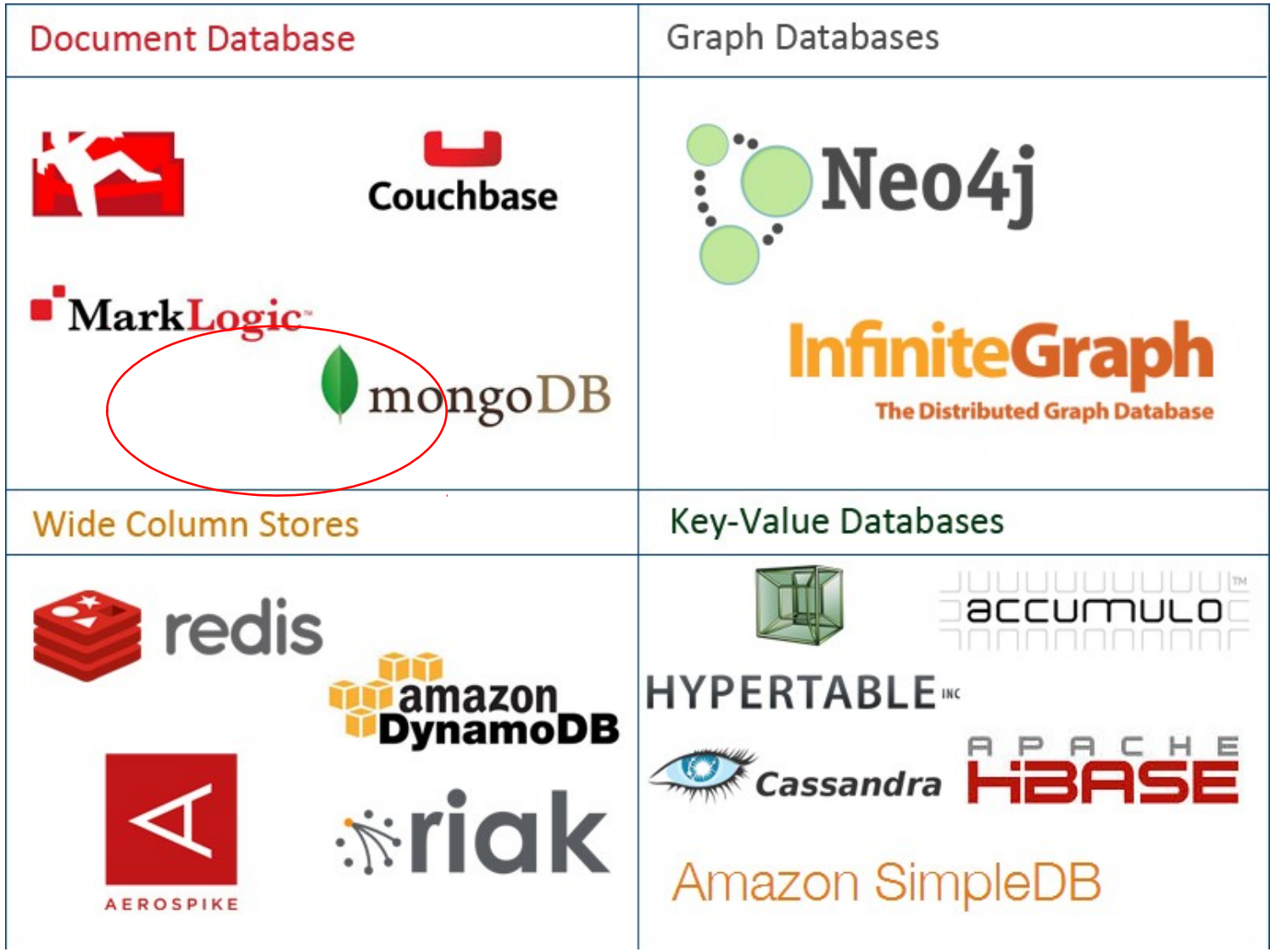
Why NoSQL

(NoSQL = Non SQL or Not only SQL)

A NoSQL database provides a mechanism for storage and retrieval o f data that is modeled in means other than the tabular relations used in relational databases.

Types

- Document
- Graph
- Key-value
- Wide-column



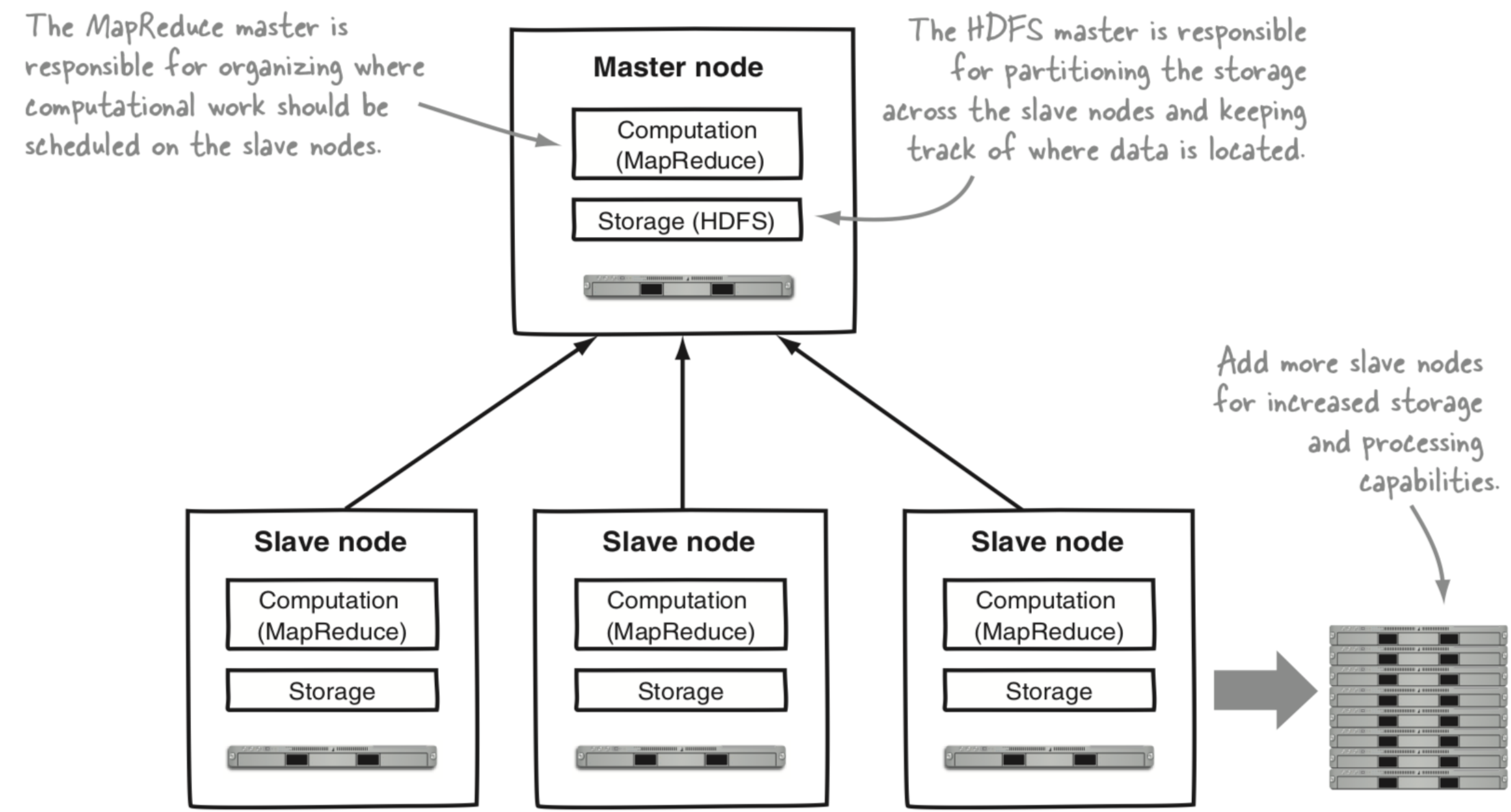
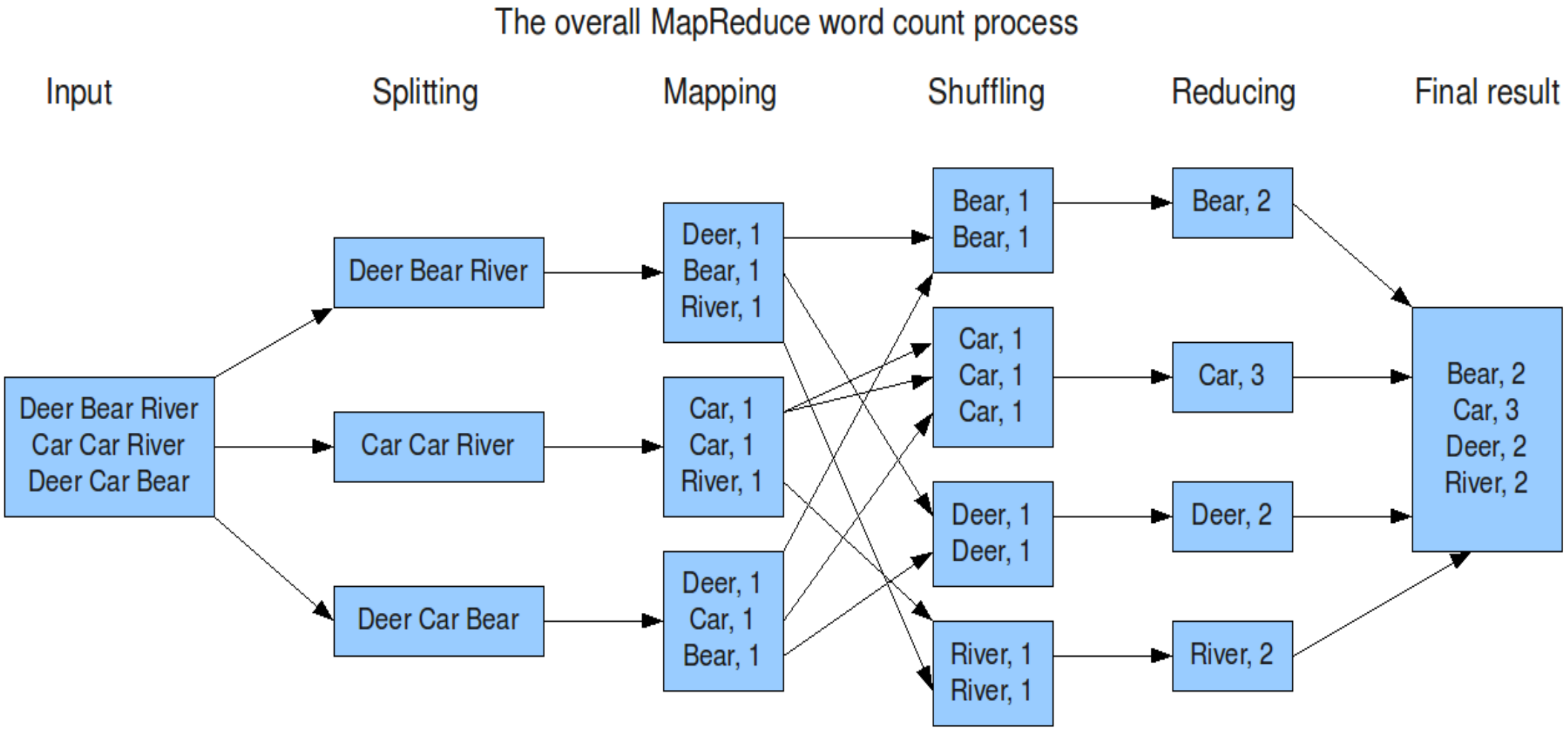
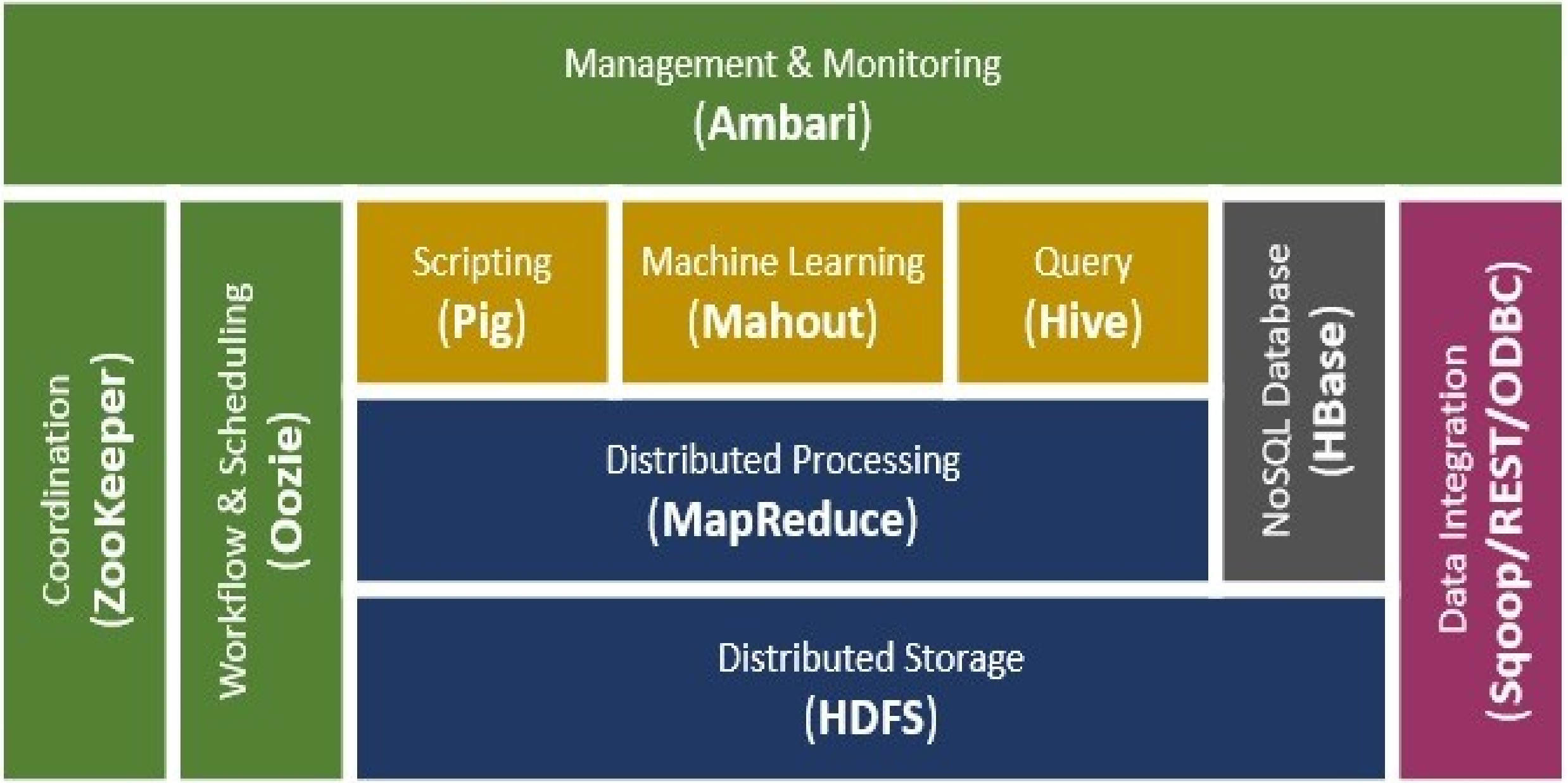


Figure 1.2 **High-level Hadoop architecture**


The overall MapReduce word count process
Input
Splitting
Mapping
Shuffling
Reducing
Final result
Deer Bear River
Car Car River
Deer Car Bear
Deer Bear River
Car Car River
Deer Car Bear
Deer, 1
Bear, 1
River, 1
Car, 1
Car, 1
River, 1
Deer, 1
Car, 1
Bear, 1
Bear, 1
Bear, 1
Car, 1
Car, 1
Car, 1
Deer, 1
Deer, 1
River, 1
River, 1
Bear, 2
Car, 3
Deer, 2
River, 2
Bear, 2
Car, 3
Deer, 2
River, 2

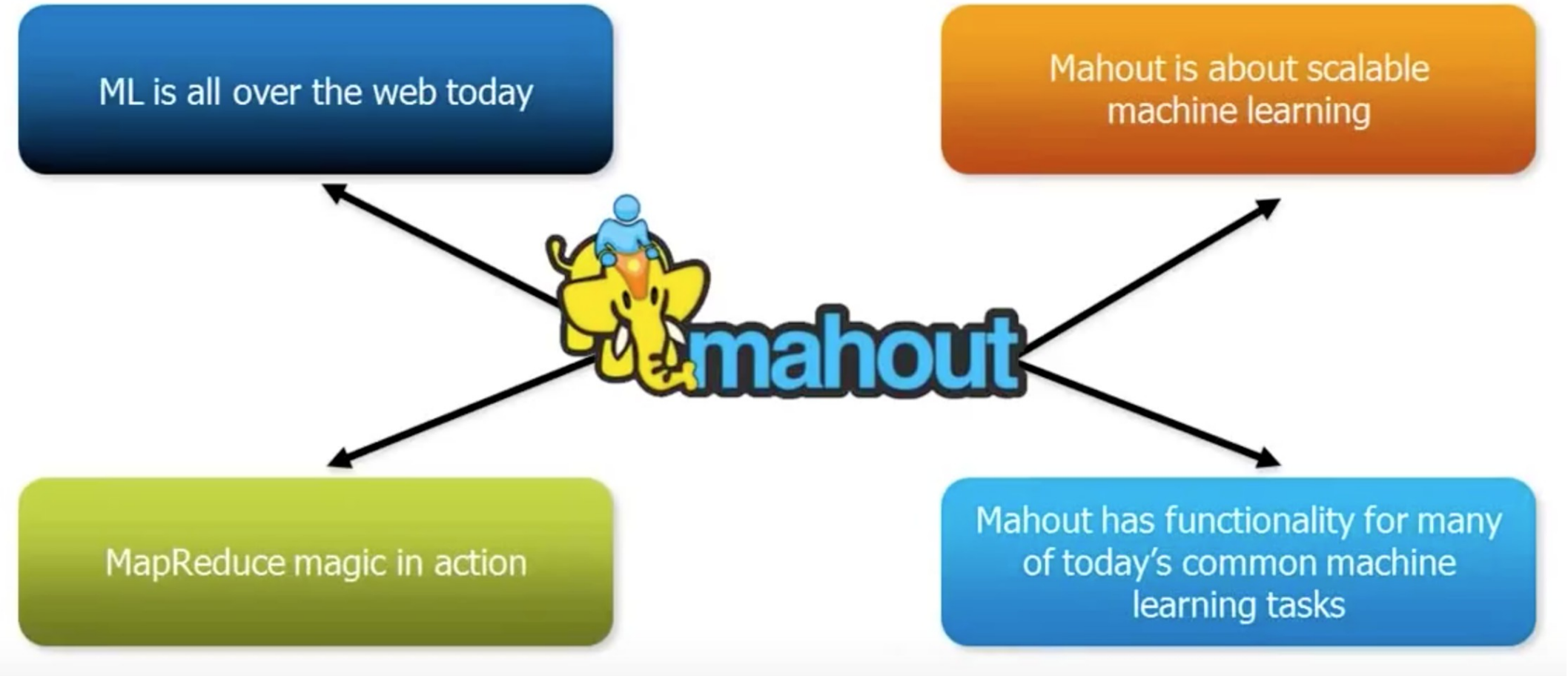
Apache Hadoop Ecosystem



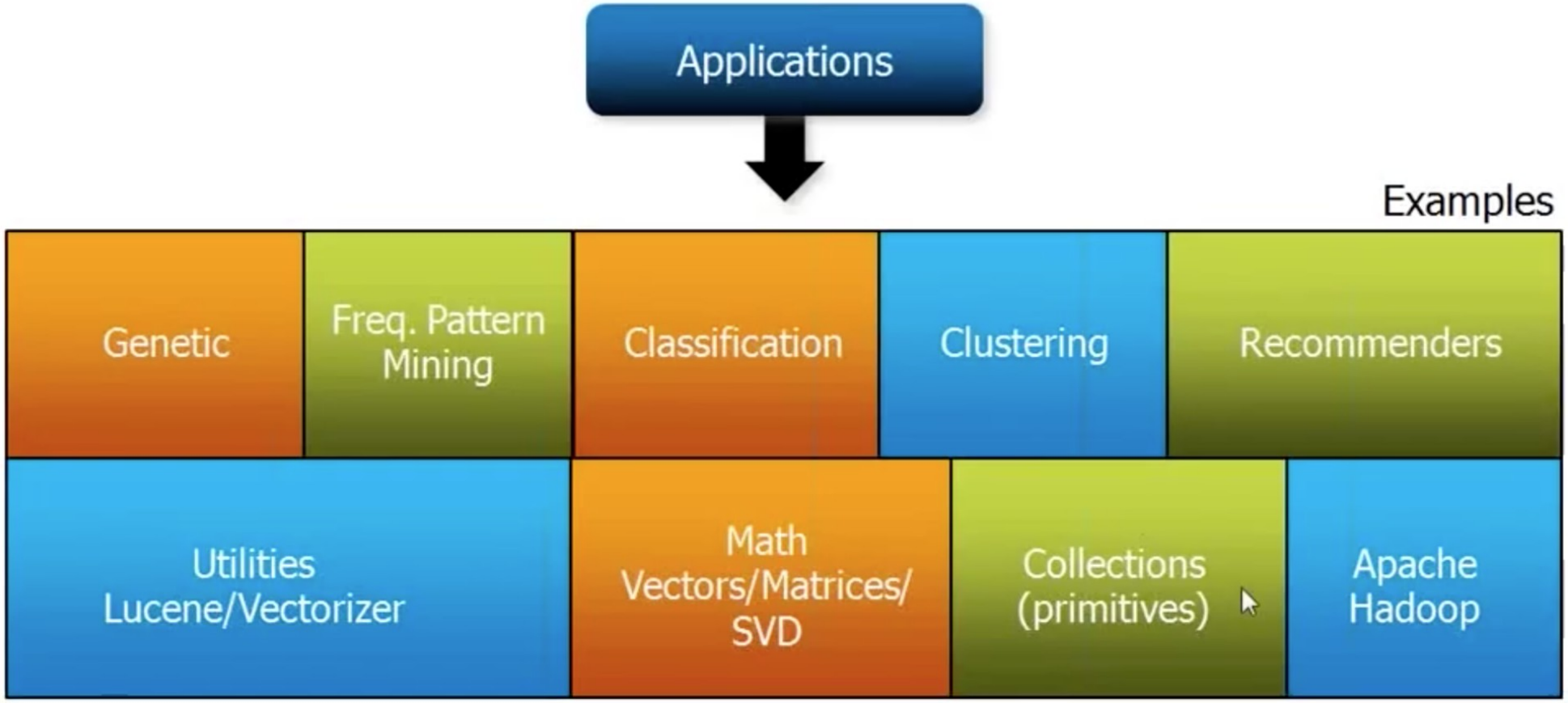
MACHINE LEARNING - TOOLS

DATA SIZE	CLASSFICATION	TOOLS
Lines Sample Data	Analysis and Visualization	Whiteboard,…
KBs - low MBs Prototype Data	Analysis and Visualization	Matlab, Octave, R, Processing,
MBs - low GBs Online Data	Analysis	NumPy, SciPy, Weka, BLAS/LAPACK
	Visualization	Flare, AmCharts, Raphael, Protovis
GBs - TBs - PBs Big Data	Analysis	**Mahout,** Giraph MLib

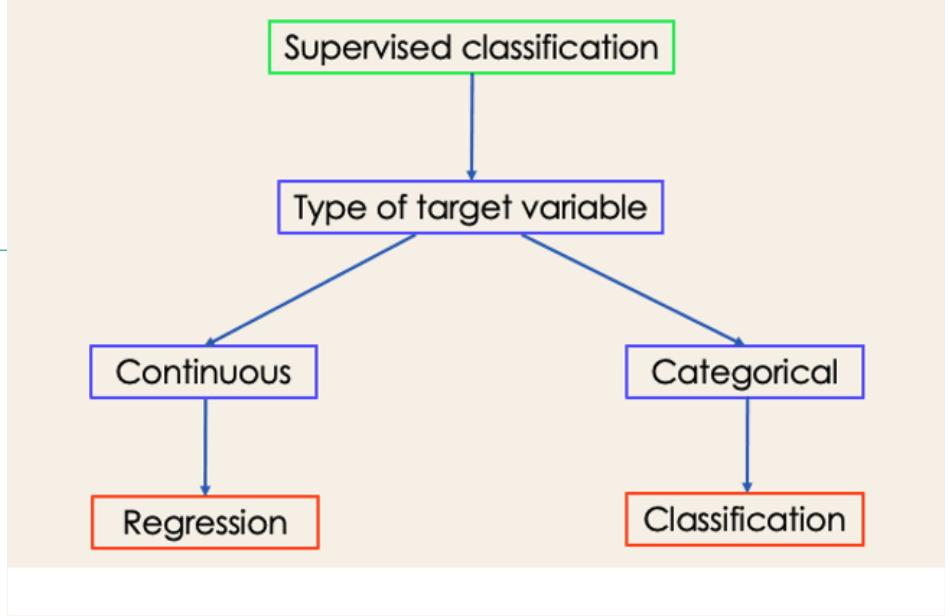
What is Mahout



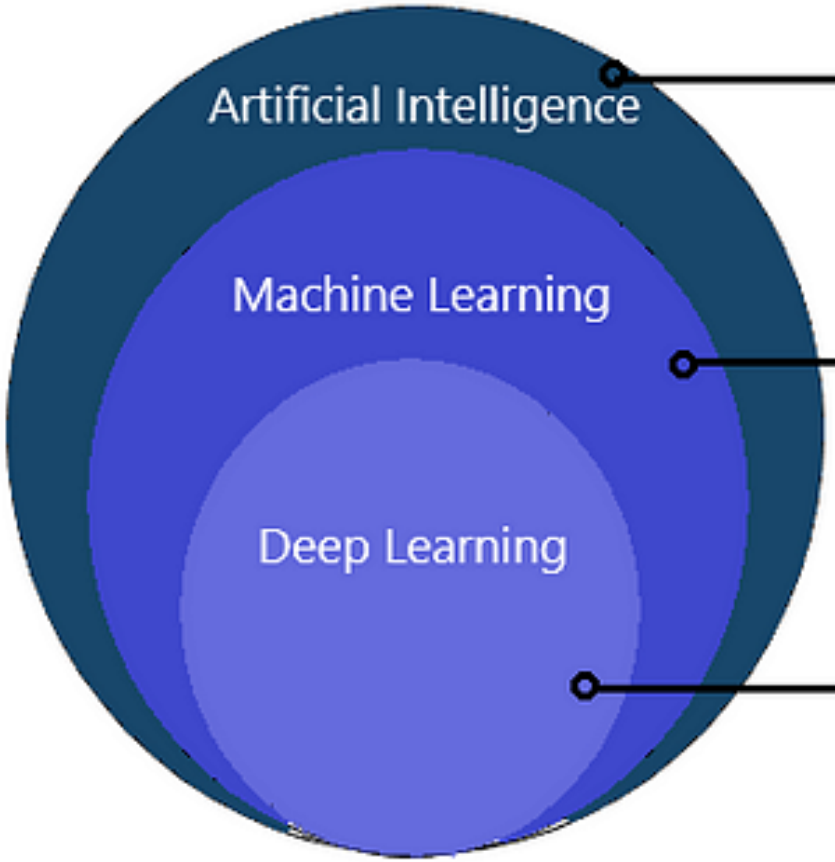
Applications of Mahout



Important terms



- Lazy Learners Vs. Eager Learners
  - Lazy: KNN
  - Eager: SVM, ANN
- Classification vs Regression
- Binary classification vs Multi-label classification
- Imbalance classification
  - Sampling Techniques: i)Cluster-based Oversampling ii) Random Undersampling, iii) SMOTE Oversampling
  - Cost-Sensitive Algorithms: i)Decision Trees, ii) Logistic Regression, iii) SVM



Artificial Intelligence

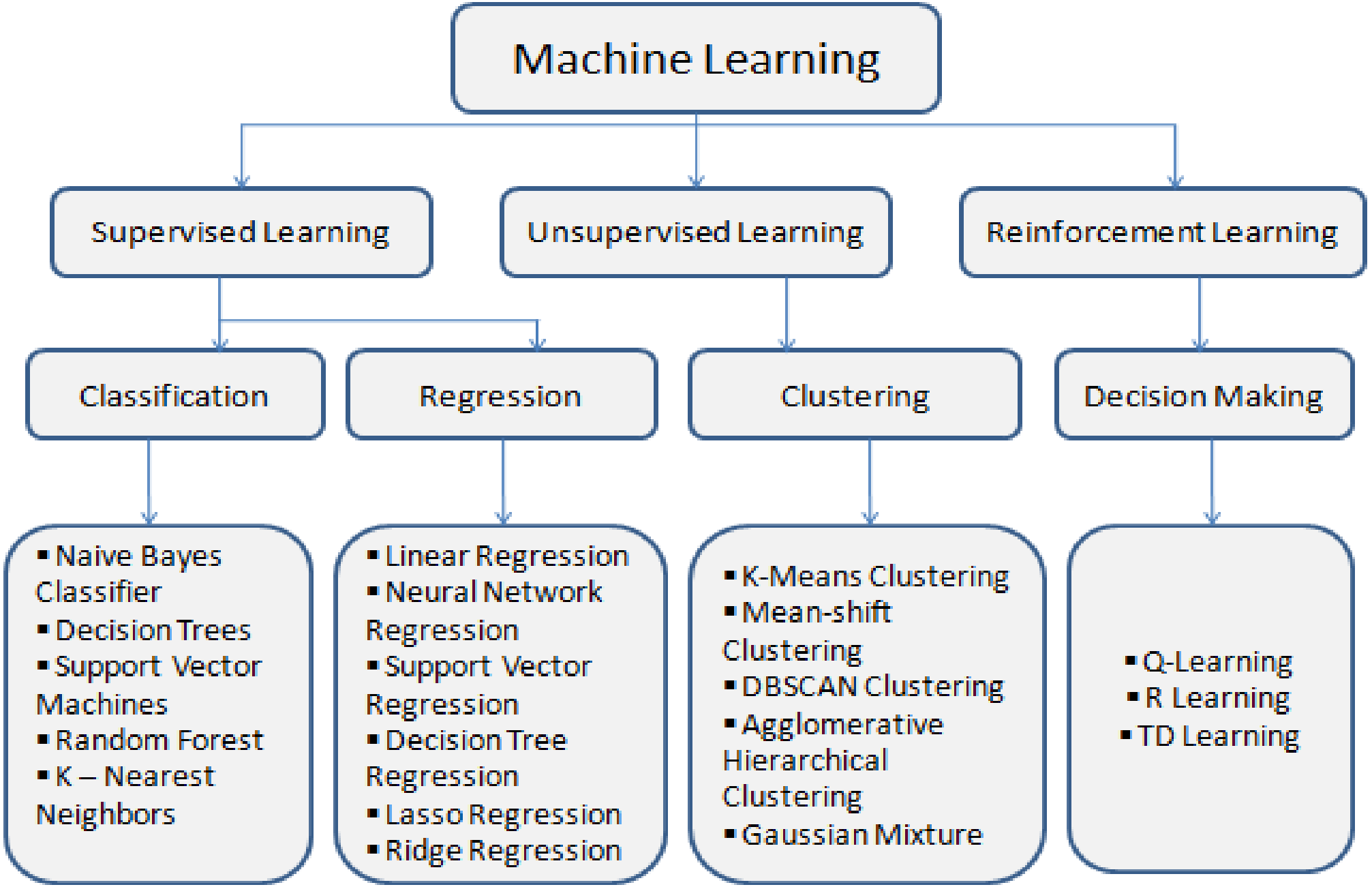
AI enables machines to think without any human intervention.

Machine Learning

Subset of AI that uses statistical learning algorithms that learn pattern in data over time

Deep Learning

Subset of ML that filters the data through multiple layers

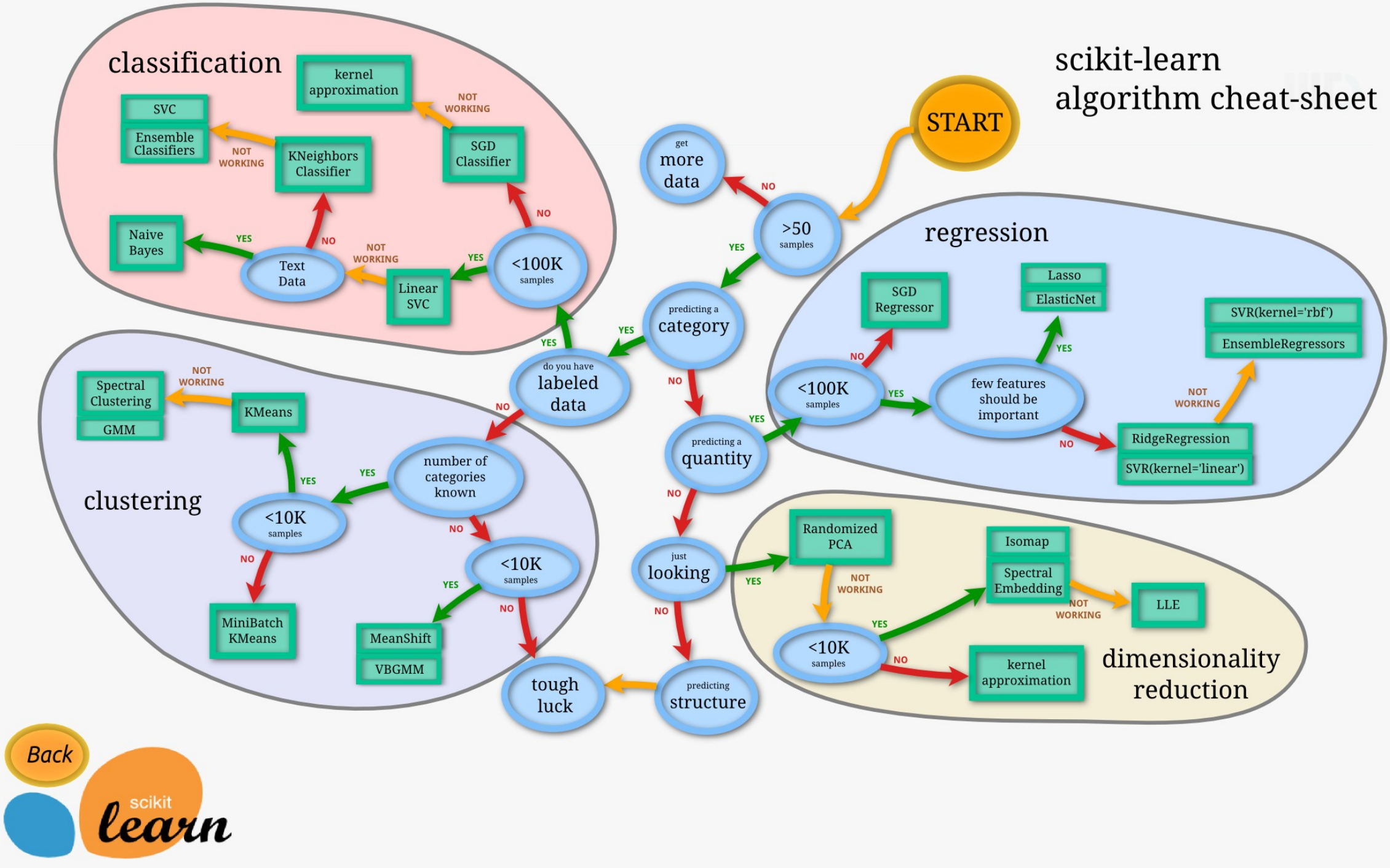

Machine Learning
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Classification
Regression
Clustering
Decision Making
▪ Naive Bayes Classifier
▪ Decision Trees
▪ Support Vector Machines
▪ Random Forest
▪ K – Nearest Neighbors
▪ Linear Regression
▪ Neural Network Regression
▪ Support Vector Regression
▪ Decision Tree Regression
▪ Lasso Regression
▪ Ridge Regression
▪ K-Means Clustering
▪ Mean-shift Clustering
▪ DBSCAN Clustering
▪ Agglomerative Hierarchical Clustering
▪ Gaussian Mixture
▪ Q-Learning
▪ R Learning
▪ TD Learning

Regression and Classification

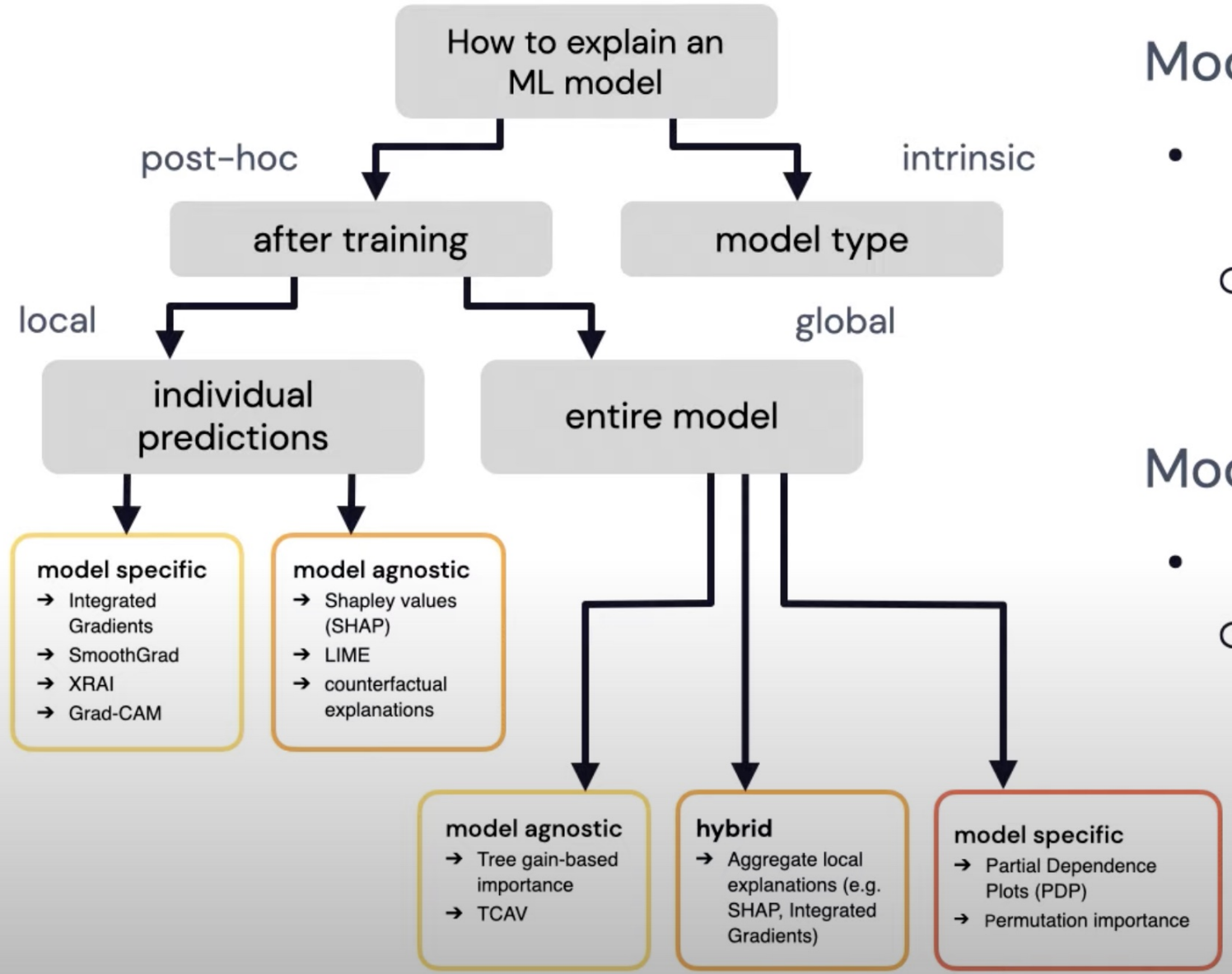
- Linear regression
- Multiple linear regression
- Classification using
  - Logistic regression (Single variable, Multiple variable)
  - Ridge Regression
  - Lasso Regression
  - Elastic Net Regression
- Machine learning techniques
  - Artificial Neural network
  - Support vector machine
  - Random Forest
  - Hidden Markov Model

Python Libraries

- General Libraries (Data science)
  - **NumPy**
  - **SciPy**
  - **Pandas**
- Visualization libraries
  - **matplotlib**
  - **Seaborn**
- Machine Learning
  - **Scikit-learn**
  - **XGBoost**
  - **NLTK (Natural Language Toolkit)**
  - **OpenCV**
  - **TensorFlow**
  - **PyTorch**


scikit-learn
algorithm cheat-sheet
START
get more data
NO
>50 samples
YES
predicting a category
YES
do you have labeled data
YES
<100K samples
NO
Linear SVC
NOT WORKING
YES
Text Data
NO
Naive Bayes
YES
KNeighbors Classifier
NOT WORKING
SVC
Ensemble Classifiers
NOT WORKING
kernel approximation
NOT WORKING
SGD Classifier
NO
classification
NO
predicting a quantity
YES
<100K samples
NO
SGD Regressor
YES
few features should be important
YES
Lasso
ElasticNet
NO
RidgeRegression
SVR(kernel='linear')
NOT WORKING
SVR(kernel='rbf')
EnsembleRegressors
regression
NO
number of categories known
YES
<10K samples
YES
KMeans
NOT WORKING
Spectral Clustering
GMM
NO
MiniBatch KMeans
NO
<10K samples
YES
MeanShift
VBGMM
NO
clustering
NO
just looking
YES
Randomized PCA
NOT WORKING
Isomap
Spectral Embedding
NOT WORKING
LLE
YES
<10K samples
NO
kernel approximation
dimensionality reduction
NO
predicting structure
tough luck
Back
scikit learn

Taxonomy: interpretable machine learning methods



Model Specific

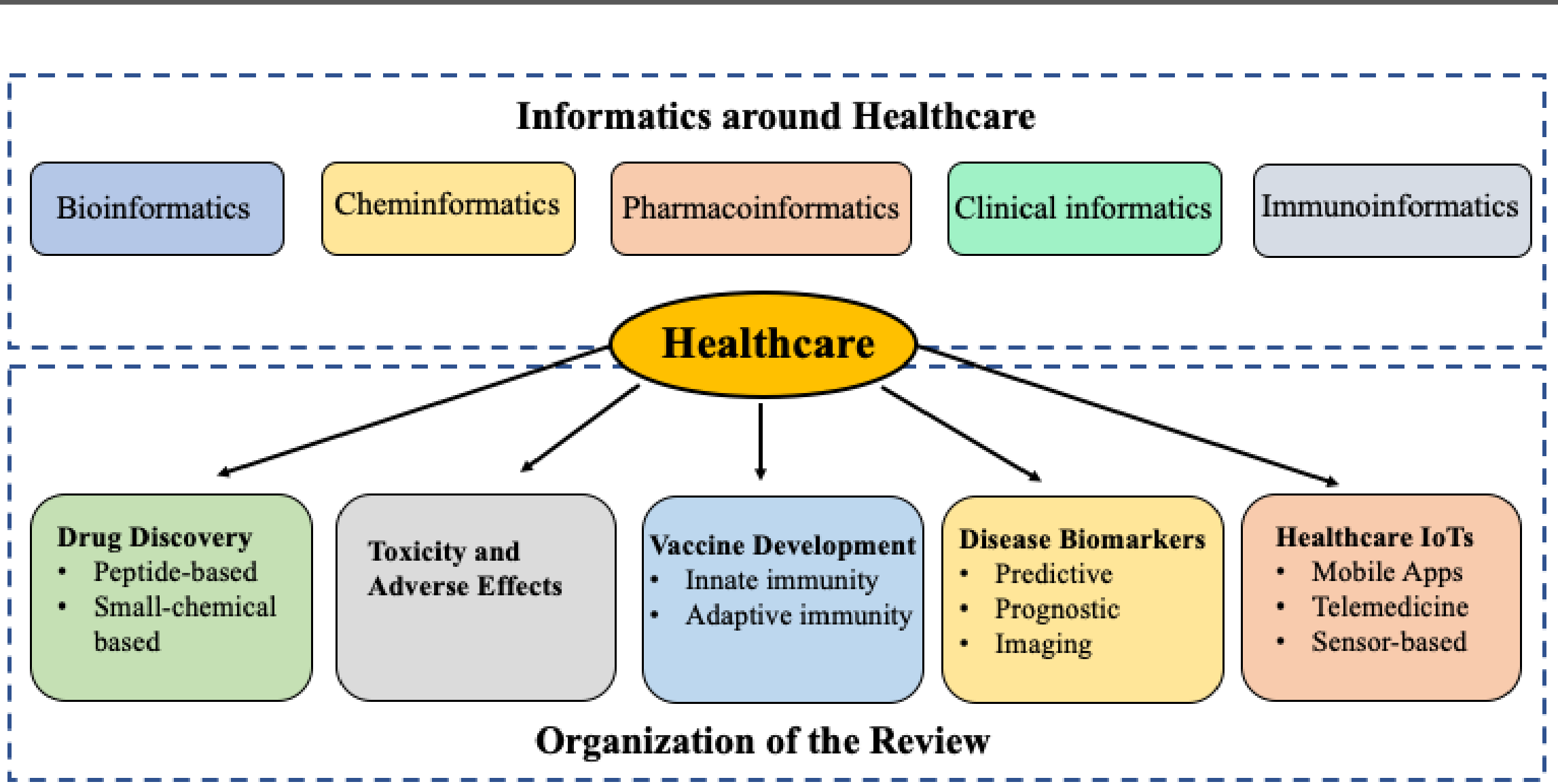
- Only works for specific models due to definition
  - e.g. neural network gradients, tree-based feature importances

Model Agnostic

- Portable across model definitions
  - e.g. Permutation feature importances, explainable surrogates

Intrinsic or Inherently Interpretable models

- **Prototype Based Models**
  - **Use prototypes or representative instances to make prediction**
  - **Example: k-Nearest Neighbors (k-NN), Exemplar-SVMs, Gaussian Mixture Models (GMM)**
- **Rule Based models**
  - **Decision Tree, Bayesian Classification, Random forest**
- **Risk Scores**
  - **Logistic Regression with Coefficients, Risk Assessment Models**
- **Generalized Additive models**
  - **Combine the flexibility of non-linear models with the transparency of additive components (e.g., Poisson, Binomial)**


Informatics around Healthcare
Bioinformatics
Cheminformatics
Pharmacoinformatics
Clinical informatics
Immunoinformatics
Healthcare
Drug Discovery
• Peptide-based
• Small-chemical based
Toxicity and Adverse Effects
Vaccine Development
• Innate immunity
• Adaptive immunity
Disease Biomarkers
• Predictive
• Prognostic
• Imaging
Healthcare IoTs
• Mobile Apps
• Telemedicine
• Sensor-based
Organization of the Review

THPdb

A database of FDA approved therapeutic peptides and proteins

Welcome to Home Page of THPdb

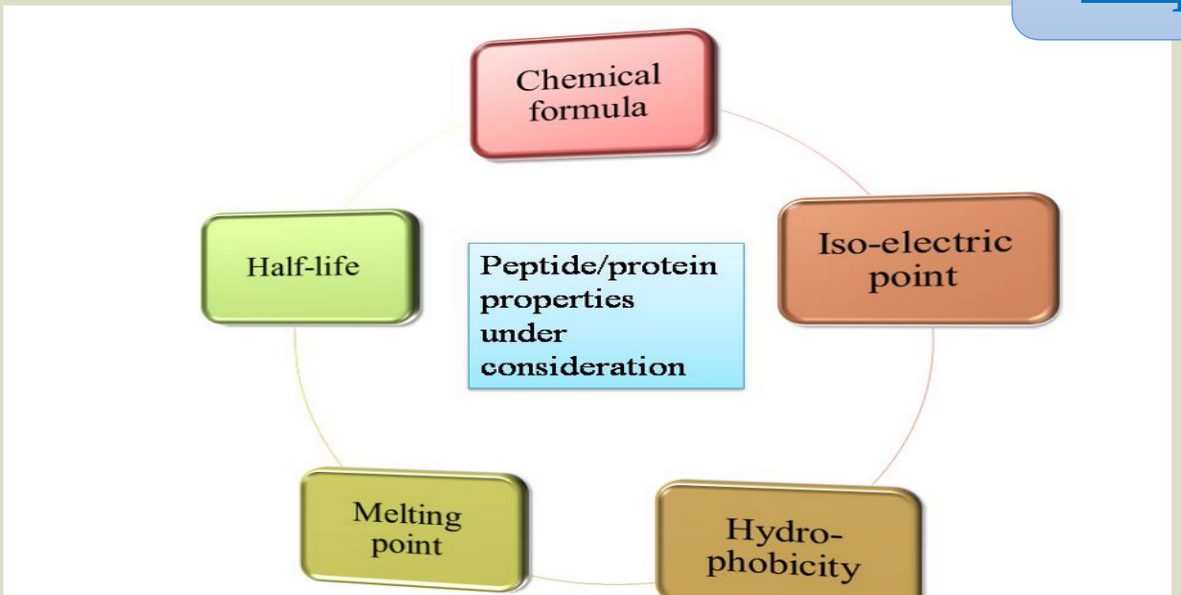
Therapeutic Proteins

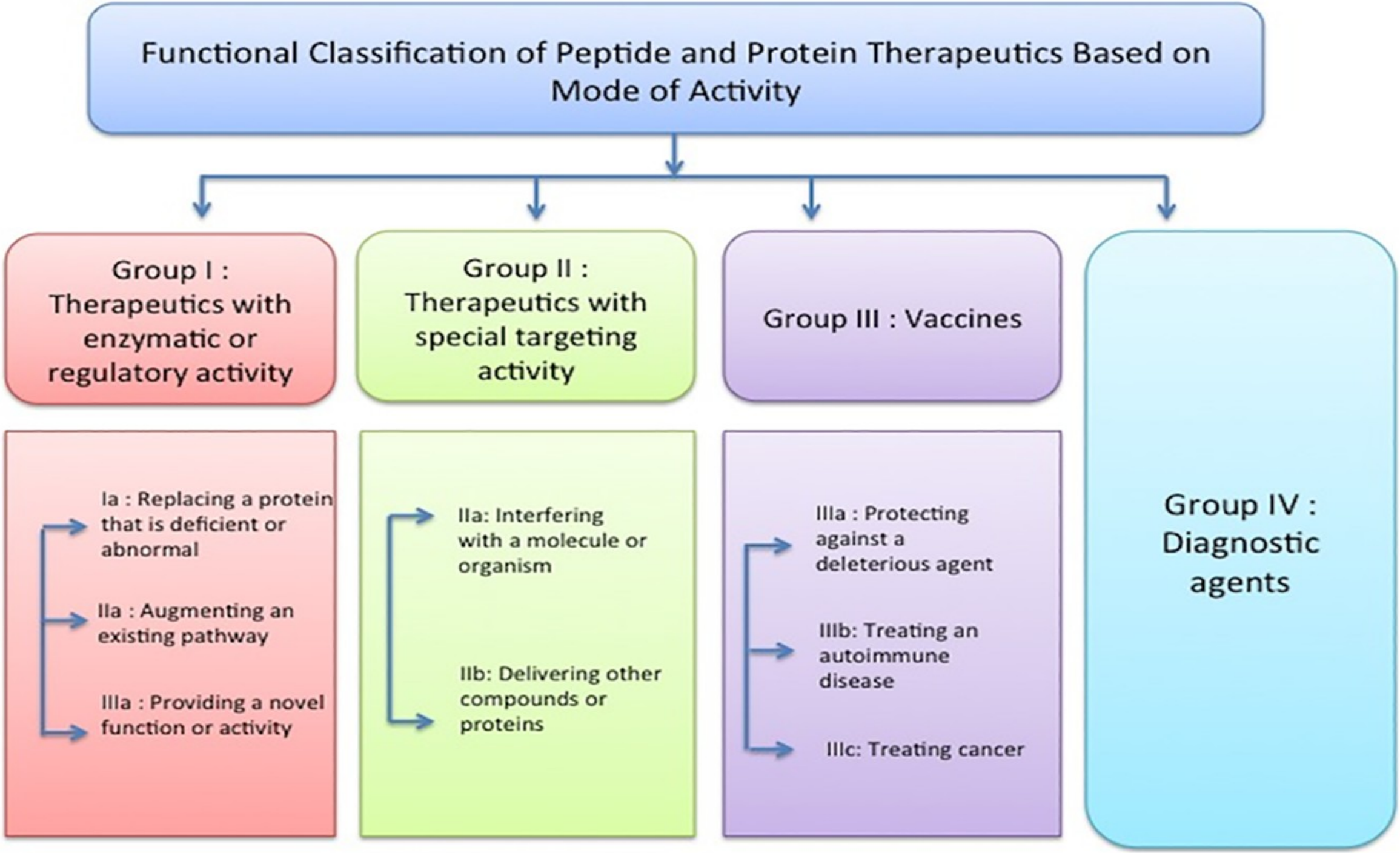
On therapeutic prespective, there is tremendous opportunity in terms of harnessing protein therapeutics to alleviate disease. Once a rarely used subset of medical treatments, protein therapeutics have increased dramatically in number and frequency of use since the introduction of the first recombinant protein therapeutic — human insulin — about 30 years ago. The pharmaceutical industry is viewing therapeutic proteins with a renewed interest. On going research is investigating a myriad of therapeutic peptides to study and improve their availability and efficacy.

What is THPdb?

THPdb is a comprehensive database based on approved and approved/investigational therapeutic peptides compiling important information about these peptides, like their description, sequence, indication, mechanism of action, pharmacodynamics, toxicity, metabolism, absorption, half life, volume of distribution, clearance rate, patent information, interaction with other drugs, targets, physicochemical properties, etc. These peptides have been classified into four categories according to their application, making it easy for the user to access them. Therapeutic peptides are modified in different ways so as to alter their properties and then sold under different brand names by various companies. THPdb provides detailed description of such brands in a user-friendly way to enable quick access of relevant information on the peptides. All information compiled in this database has been extracted from research papers, patents, pharmaceutical company websites catering to product details, drugba

https://webs.iiitd.edu.in/raghava/thpdb/




Functional Classification of Peptide and Protein Therapeutics Based on Mode of Activity
Group I : Therapeutics with enzymatic or regulatory activity
Group II : Therapeutics with special targeting activity
Group III : Vaccines
Group IV : Diagnostic agents
Ia : Replacing a protein that is deficient or abnormal
IIa : Augmenting an existing pathway
IIIa : Providing a novel function or activity
IIa: Interfering with a molecule or organism
IIb: Delivering other compounds or proteins
IIIa : Protecting against a deleterious agent
IIIb: Treating an autoimmune disease
IIIc: Treating cancer

Drug Delivery

crdd.osdd.net/raghava/cellppd/

Google State Bank of India Home - PubMed -

CellPPD: Designing of Cell Penetrating Peptides

Home | Design Peptide | Multiple Peptides | Protein Scaning | Motif Scaning | Motif List | Major Features | Algorithm | Help | Datasets

Welcome to CellPPD

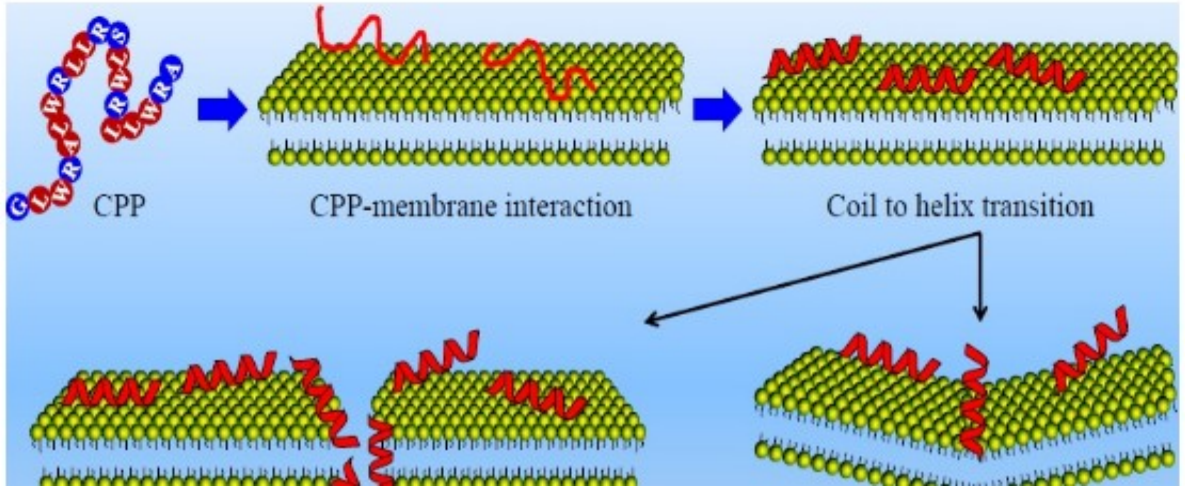
utam *et al.*: *In silico* approaches for designing highly effective cell penetrating peptides. *Journal of Translational Medicine* 3 11:74.Link

CellPPD is an *in silico* method, which is developed to predict and design efficient cell penetrating peptides (CPPs). The main dataset used in this method consists of 708 experimentally validated CPPs.

Major Features include:

(1) **Desing Peptide:** This module allows user to generate all possible single mutant analogues of their peptides and predict whether the analogue is cell penetrating or not.

(2) **Multiple Peptides:** This module of CellPPD allows user to predict number of CPPs in peptides submitted by the

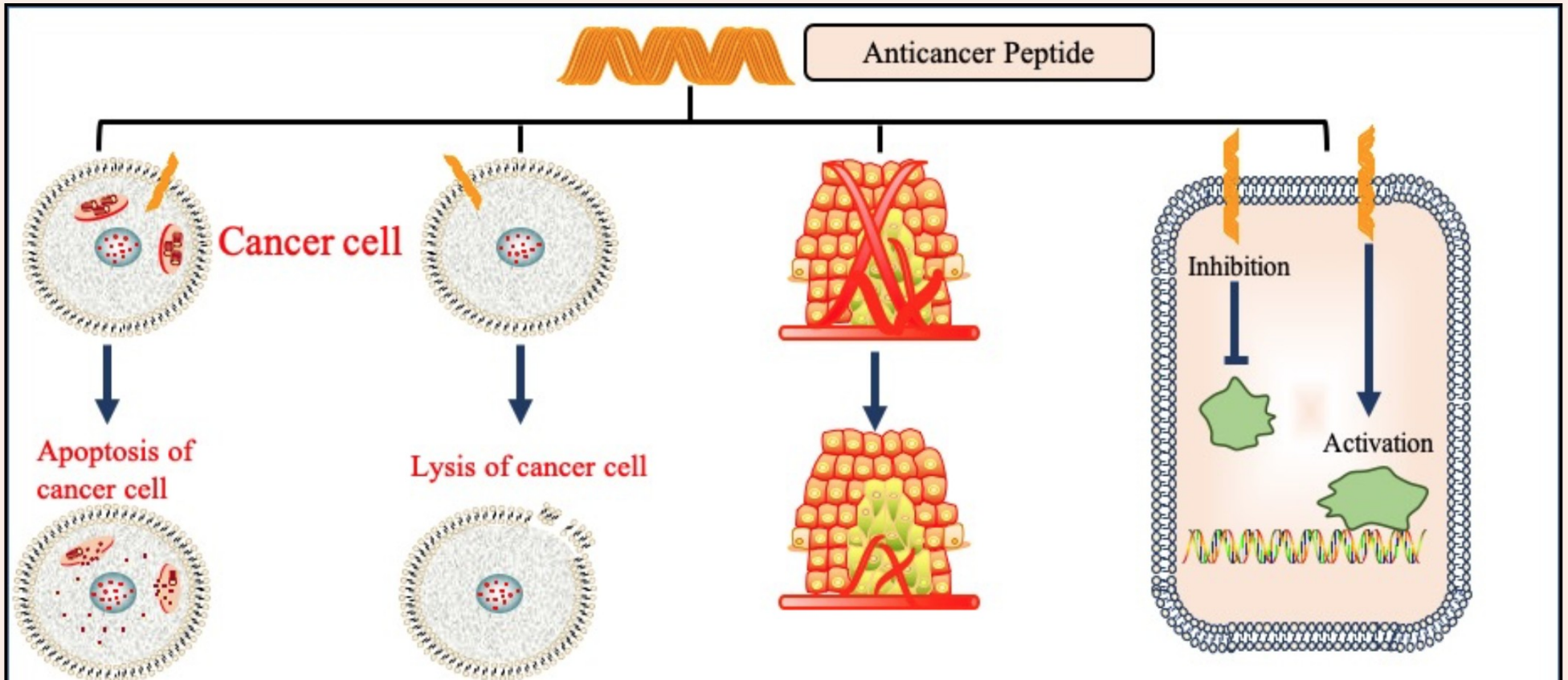


Peptide based Inhibitors

AntiCP 2.0

Welcome To AntiCP 2.0

AntiCP 2.0 is an updated version of AntiCP, developed to predict and design anticancer peptides with high accuracy. This study utilize largest possible dataset of anticancer and non-anticancer peptides. Main dataset consists of experimentally validated 861 anticancer peptides and 861 non-anticancer or validated antimicrobial peptides. Alternate dataset comprises of 970 anti-cancer peptides and 970 non-anticancer peptides (randomly pickup from Swiss-Prot).



Peptide based Inhibitors

AHTpin

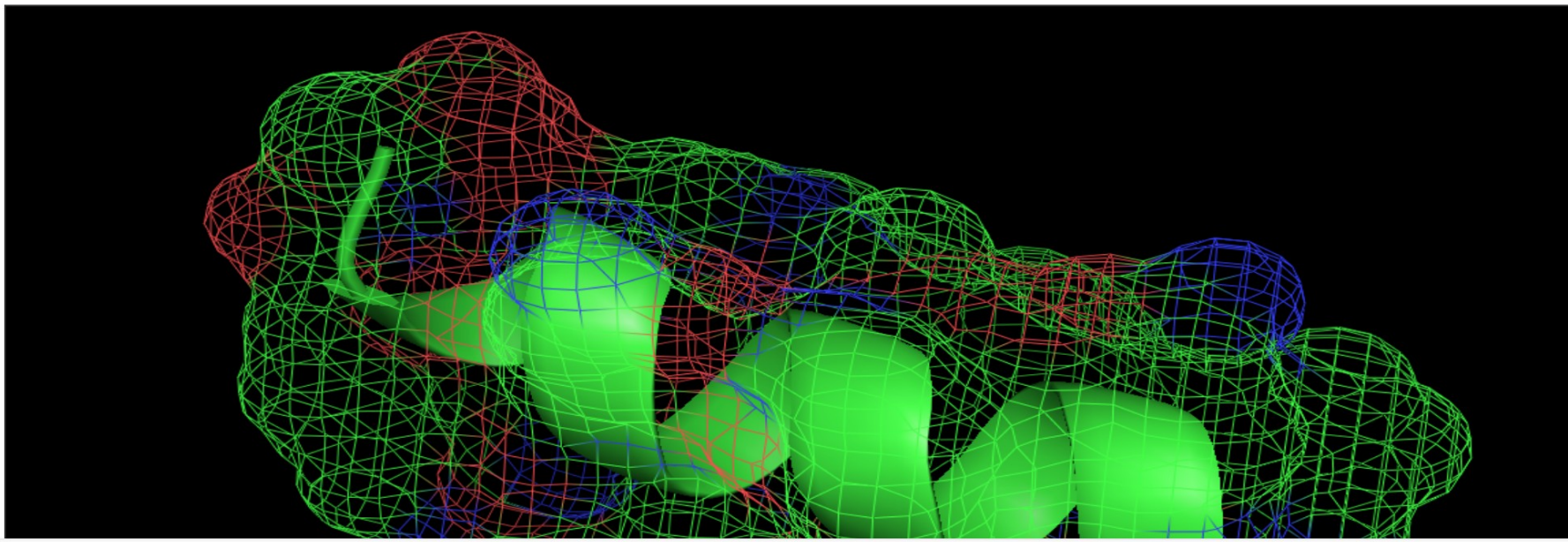
ANTIHYPERTENSIVE PEPTIDE INHIBITORS

HOME ALGORITHM DATASETS HELP TEAM CONTACT

designing of antihypertensive peptides. Sci. Rep. 5, 12512.

- Dipeptide
- Tripeptide
- Tetrapeptide
- Pentapeptide
- Hexapeptide
- 7-12 residues

Welcome to Home Page of AHTpin



Structure of Chemically Modified Peptides

- Home PEPstrMOD
- Natural Peptides
- D Amino Acids
- Terminal Modification
- Peptide Cyclization
  - N-C Cyclization
  - Disulfide Bridge (S-S)
  - Structure Modification
- Non-Natural Residue
- PTMs of Residue
- Advance Modification
- Structure Simulations
- Download
- General

Welcome to Peptide Cyclization Module for N-C cyclization

This page is designed to predict the peptide structure with N-to-C terminal cyclization. The peptide is made cyclic by incorporating a bond between Nitrogen atom of N-terminal residue and Carbon atom of C-terminal residue. For more information click help

Peptide Sequence Submission Form

Peptide sequence in plain text format Example Sequence

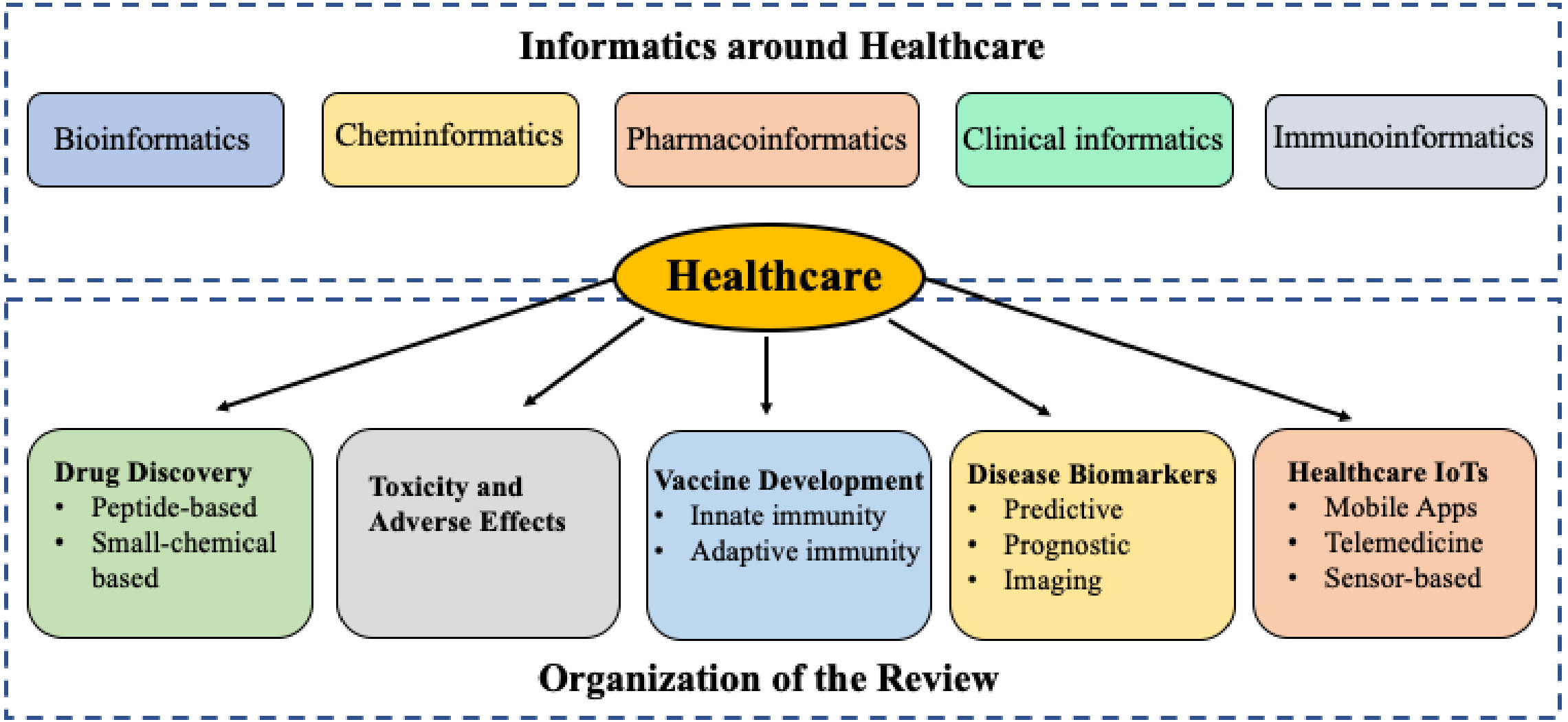
Email Address:

Advanced Options: CLICK

Reset or clear form Submit sequence for prediction

Peptide based therapeutics

Tool	Description (Link)
AniAMPpred	Prediction of antimicrobial peptides in animal kingdom (https://aniamppred.anvil.app/)
B3Pdb	Compilation of Blood Brain Barrier Penetrating Peptides (https://webs.iiitd.edu.in/raghava/b3pdb/)
B3Pred	Blood-Brain Barrier penetrating peptides prediction (https://webs.iiitd.edu.in/raghava/b3pred/)
AlgPred 2.0	Highly accurate method for predicting allergic proteins (https://webs.iiitd.edu.in/raghava/algpred2 /)
AntiCP 2.0	Improved method for identification of anticancer peptides (https://webs.iiitd.edu.in/raghava/anticp2/)
HemoPI-MOD	Hemolytic potency of chemically modified peptides (https://webs.iiitd.edu.in/raghava/hemopimod/)
Antifp	Prediction of antifungal peptides (https://webs.iiitd.edu.in/raghava/antifp/)
AntiMPmod	Antimicrobial potential of chemically modified peptides (https://webs.iiitd.edu.in/raghava/antimpmod/)
AntiTbPred	Prediction of antitubercular peptides (https://webs.iiitd.edu.in/raghava/antitbpred/)
CellPPD-MOD	Computation of chemically modified cell penetrating peptides (https://webs.iiitd.edu.in/raghava/cellppdmod/)
PlifePred	Estimation of half-life of peptides in blood (https://webs.iiitd.edu.in/raghava/plifepred/)
TopicalPdb	Repository of topically delivered peptides (https://webs.iiitd.edu.in/raghava/topicalpdb/)
THPdb	Compilation of peptide/protein based therapeutic molecules (https://webs.iiitd.edu.in/raghava/thpdb/)
CPPSite2	Database of cell-penetrating peptides (https://webs.iiitd.edu.in/raghava/cppsite/)
AHTpin	Designing antihypertensive peptides (https://webs.iiitd.edu.in/raghava/ahtpin/)
AntiAngioPred	Prediction of anti-angiogenic peptides (http://clri.res.in/subramanian/tools/antiangiopred/)
CancerPPD	Database of anticancer peptides and proteins (https://webs.iiitd.edu.in/raghava/cancerppd/)
TumorHPD	Designing tumor homing peptides (https://webs.iiitd.edu.in/raghava/tumorhpd/)
AntiBP2	Webserver for antibacterial peptide prediction(https://webs.iiitd.edu.in/raghava/antibp2/)


Informatics around Healthcare
Bioinformatics
Cheminformatics
Pharmacoinformatics
Clinical informatics
Immunoinformatics
Healthcare
Drug Discovery
• Peptide-based
• Small-chemical based
Toxicity and Adverse Effects
Vaccine Development
• Innate immunity
• Adaptive immunity
Disease Biomarkers
• Predictive
• Prognostic
• Imaging
Healthcare IoTs
• Mobile Apps
• Telemedicine
• Sensor-based
Organization of the Review

Current Topics in Medicinal Chemistry, 2013, *13*, 1172-1191

Open Source Software and Web Services for Designing Therapeutic Molecules

Deepak Singla[1,2], Sandeep Kumar Dhanda[1], Jagat Singh Chauhan[1], Anshu Bhardwaj[3], Samir K. Brahmachari[3,4], Open Source Drug Discovery Consortium[3] and Gajendra P.S. Raghava[1,*]

[1]*Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh, India;* [2]*Centre for Microbial Biotechnology, Panjab University, Chandigarh, India;* [3]*CSIR-Open Source Drug Discovery Unit, New Delhi, India;* [4]*CSIR-Institute of Genomics and Integrative Biology, New Delhi, India*

Abstract: Despite the tremendous progress in the field of drug designing, discovering a new drug molecule is still a challenging task. Drug discovery and development is a costly, time consuming and complex process that requires millions of dollar and 10-15 years to bring new drug molecules in the market. This huge investment and long-term process are attributed to high failure rate, complexity of the problem and strict regulatory rules, in addition to other factors. Given the availability of 'big' data with ever improving computing power, it is now possible to model systems which is expected to provide time and cost effectiveness to drug discovery process. Computer Aided Drug Designing (CADD) has emerged as a fast alternative method to bring down the cost involved in discovering a new drug. In past, numerous computer programs have been developed across the globe to assist the researchers working in the field of drug discovery. Broadly, these programs can be classified in three categories, freeware, shareware and commercial software. In this review, we have described freeware or open-source software that are commonly used for designing therapeutic molecules. Major emphasis will be on software and web services in the field of chemo- or pharmaco-informatics that includes *in silico* tools used for computing molecular descriptors, inhibitors designing against drug targets, building QSAR models, and ADMET properties.

REVIEW ARTICLE

Overview of Free Software Developed for Designing Drugs Based on Protein-Small Molecules Interaction

Piyush Agrawal[1,2], Pawan Kumar Raghav[1], Sherry Bhalla[1], Neelam Sharma[1], and Gajendra P.S. Raghava[1,2,*]

[1]*Center for Computational Biology, Indraprastha Institute of Information Technology, New Delhi, India;* [2]*CSIR-Institute of Microbial Technology, Chandigarh, India*

ARTICLE HISTORY

Received: April 05, 2018
Revised: June 08, 2018
Accepted: August 08, 2018

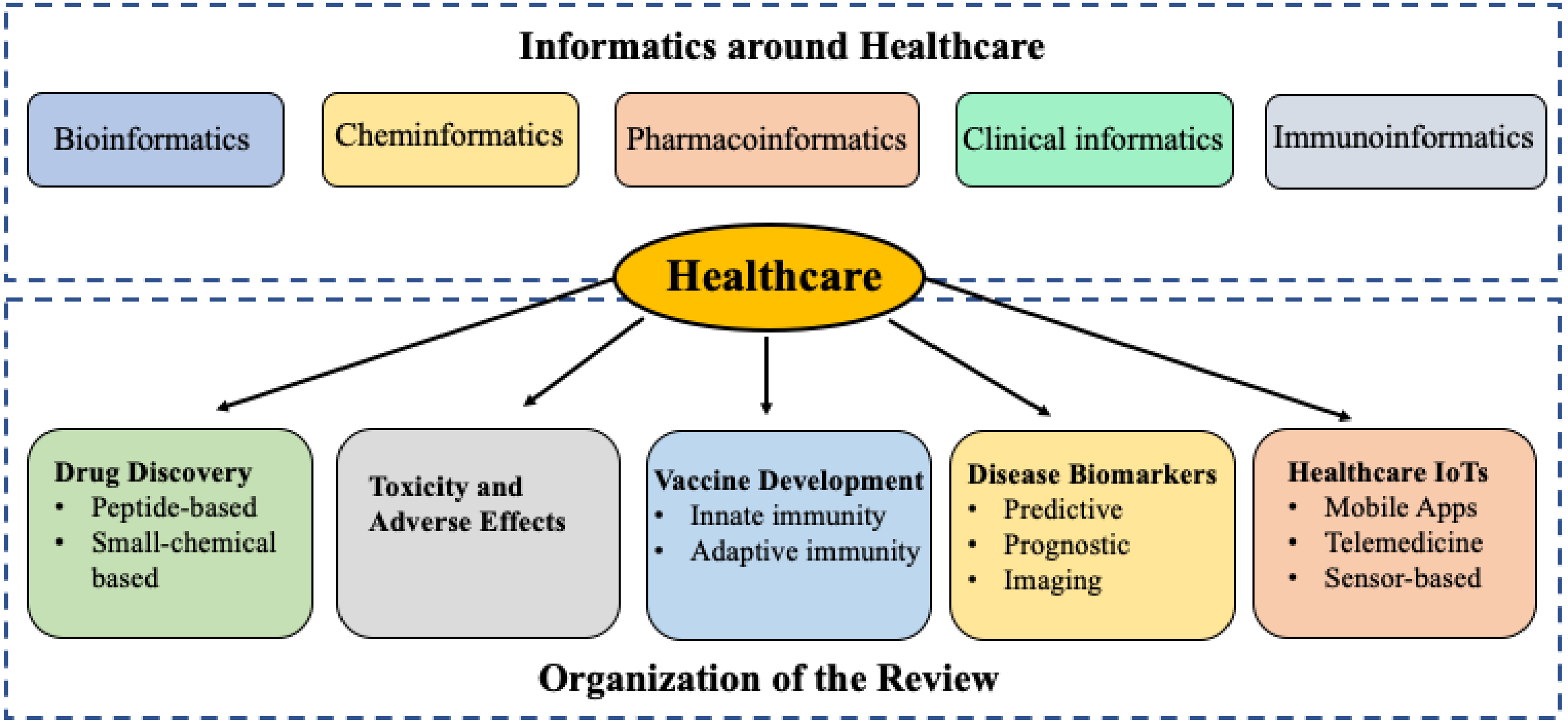
DOI:
10.2174/1568026618666180816155131

Abstract: One of the fundamental challenges in designing drug molecule against a disease target or protein is to predict binding affinity between target and drug or small molecule. In this review, our focus will be on advancement in the field of protein-small molecule interaction. This review has been divided into four major sections. In the first section, we will cover software developed for protein structure prediction. This will include prediction of binding pockets and post-translation modifications in proteins. In the second section, we will discuss software packages developed for predicting small-molecule interacting residues in a protein. Advances in the field of docking particularly advancement in the knowledge-based force fields will be discussed in the third part of the review. This section will also cover the method developed for predicting affinity between protein and drug molecules. The fourth section of the review will describe miscellaneous techniques used for designing drug molecules, like pharmacophore modelling. Our major emphasis in this review will be on computational tools that are available free for academic use

Keywords: Protein-small molecule interaction, Structure Prediction, Docking, Pharmacophore, Molecular Dynamics, Post Translational Modifications.

Small Chemical-based Drug Discovery

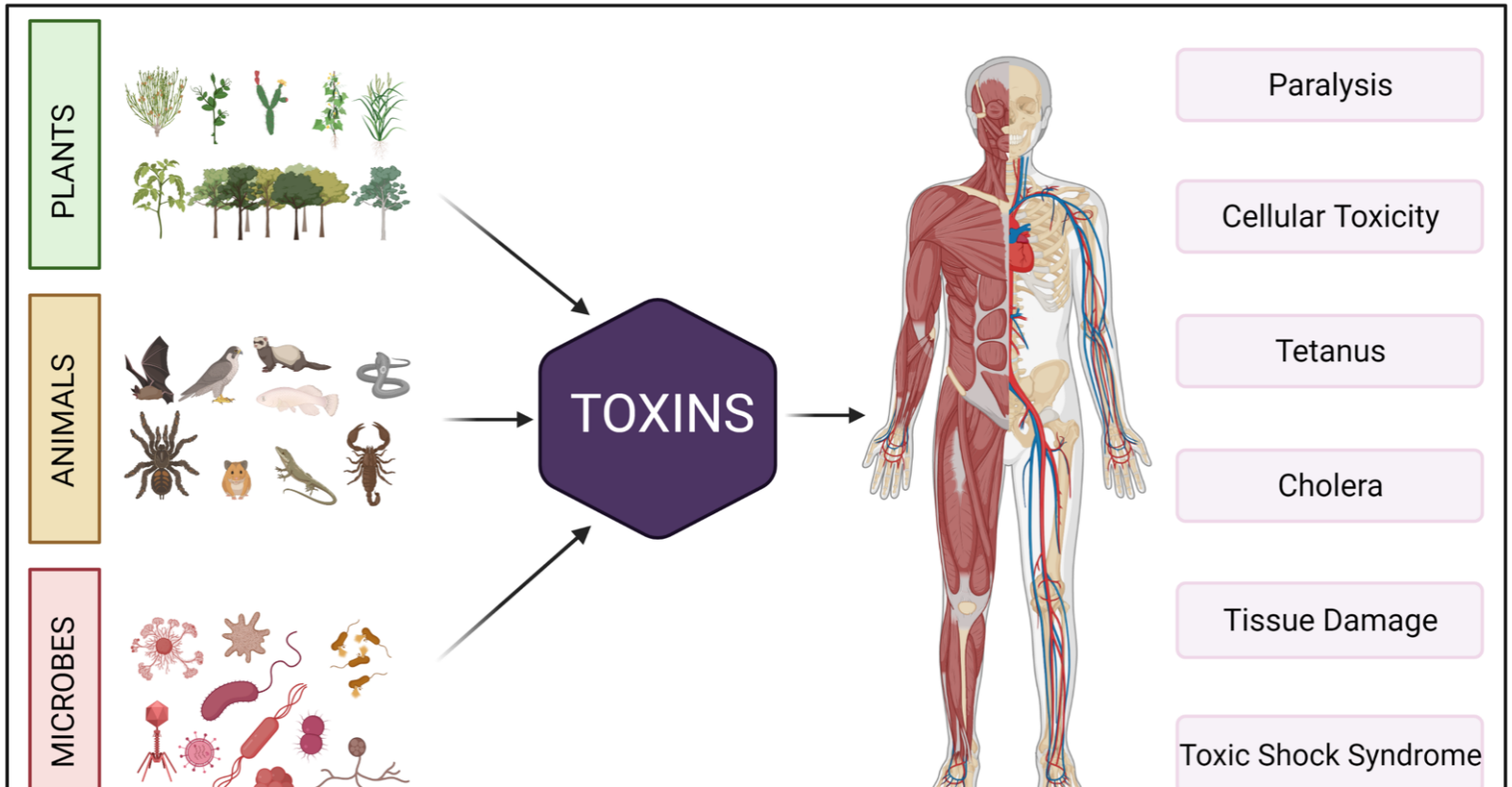
Tool	Description (Link)
PubChem	Database of bioassays, compounds and substances (https://pubchem.ncbi.nlm.nih.gov/)
ChEMBL	Database of drug like molecules (https://www.ebi.ac.uk/chembldb)
Zinc15	Database of commercially-available compounds for virtual screening (http://zinc.docking.org/)
DrugBank	Comprehensive information about drugs (https://go.drugbank.com/)
BindingDB	Binding affinity of PDB ligands (http://www.bindingdb.org/)
SuperDRUG2	Database of approved/marketed drugs (http://cheminfo.charite.de/superdrug2/)
PaDEL	1D, 2D, 3D and fingerprints calculation (http://padel.nus.edu.sg/software/padeldescriptor)
CDK	Chemistry Development Kit (http://cdk.sourceforge.net)
Mordred	Molecular descriptor calculator (https://github.com/mordred-descriptor/mordred)
Dock 6	Standalone software for molecular docking (http://dock.compbio.ucsf.edu/)
AutoDock Vina	Program for molecular docking and virtual screening (http://vina.scripps.edu/)
Autodock	Molecular modelling simulation software (http://autodock.scripps.edu/)
QSAR-Co	Classification-based QSAR model development (https://sites.google.com/view/qsar-co)
DPubChem	Web tool for QSAR modeling and high-throughput virtual screening (https://www.cbrc.kaust.edu.sa/dpubchem/)
Weka	Collection of machine learning algorithm for the development of QSAR based models (https://www.cs.waikato.ac.nz/~ml/weka/)
TINKER	Software tools for molecular design (http://dasher.wustl.edu/tinker/))
Frog	Generation of free online drug conformation (http://bioserv.rpbs.jussieu.fr/cgi-bin/Frog)
Openbabel	The open source chemistry toolbox (http://openbabel.org/)
EGFRpred	QSAR based model for predicting EGFR inhibitors (https://webs.iiitd.edu.in/oscadd/egfrpred/)
MDRIpred	Prediction of inhibitor against drug tolerant M. tuberculosis (https://webs.iiitd.edu.in/oscadd/mdri/)
GDoQ	Prediction of inhibitors against Mycobacterial GlmU (https://webs.iiitd.edu.in/raghava/gdoq/)


Informatics around Healthcare
Bioinformatics
Cheminformatics
Pharmacoinformatics
Clinical informatics
Immunoinformatics
Healthcare
Drug Discovery
• Peptide-based
• Small-chemical based
Toxicity and Adverse Effects
Vaccine Development
• Innate immunity
• Adaptive immunity
Disease Biomarkers
• Predictive
• Prognostic
• Imaging
Healthcare IoTs
• Mobile Apps
• Telemedicine
• Sensor-based
Organization of the Review

Welcome to ToxinPred2

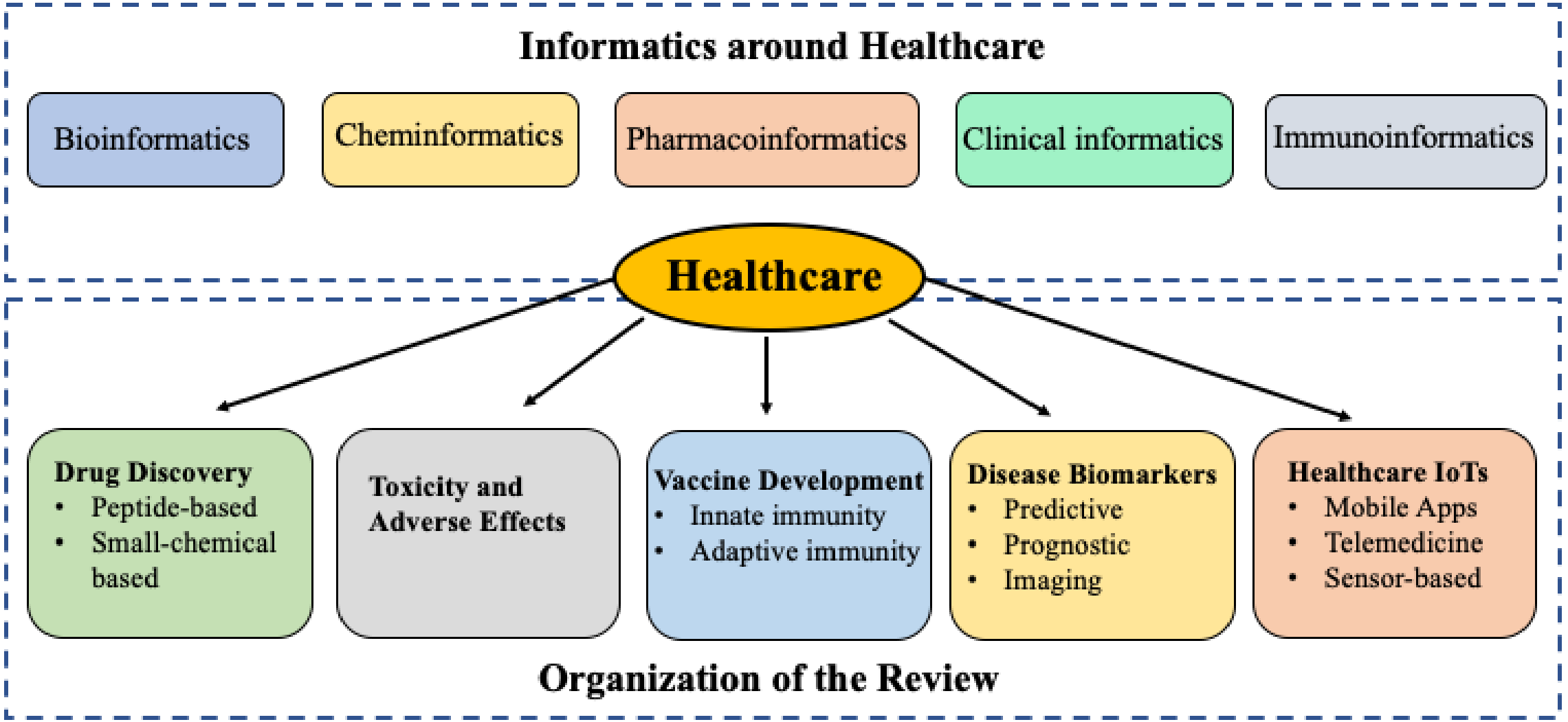
A webserver for predicting toxic and non-toxic proteins

ToxinPred2 is developed for predicting toxicity of proteins. In case user is interested in predicting toxicity of peptides then users should use our old server **ToxinPred**, which is specifically designed for peptides (See PLoS One. 2013 Sep 13;8(9):e73957). It utilizes wide range of information and techniques for prediction that includes machine learning techniques, BLAST, and MERCI. Our models have been trained on large dataset that contains 8233 toxic and 8233 non-toxic. In addition to following major modules, the server also has following important links; i) Download (Standalone Version & Datasets), ii) Algorithm, iii) Help page, iv) Team and v) Contact.

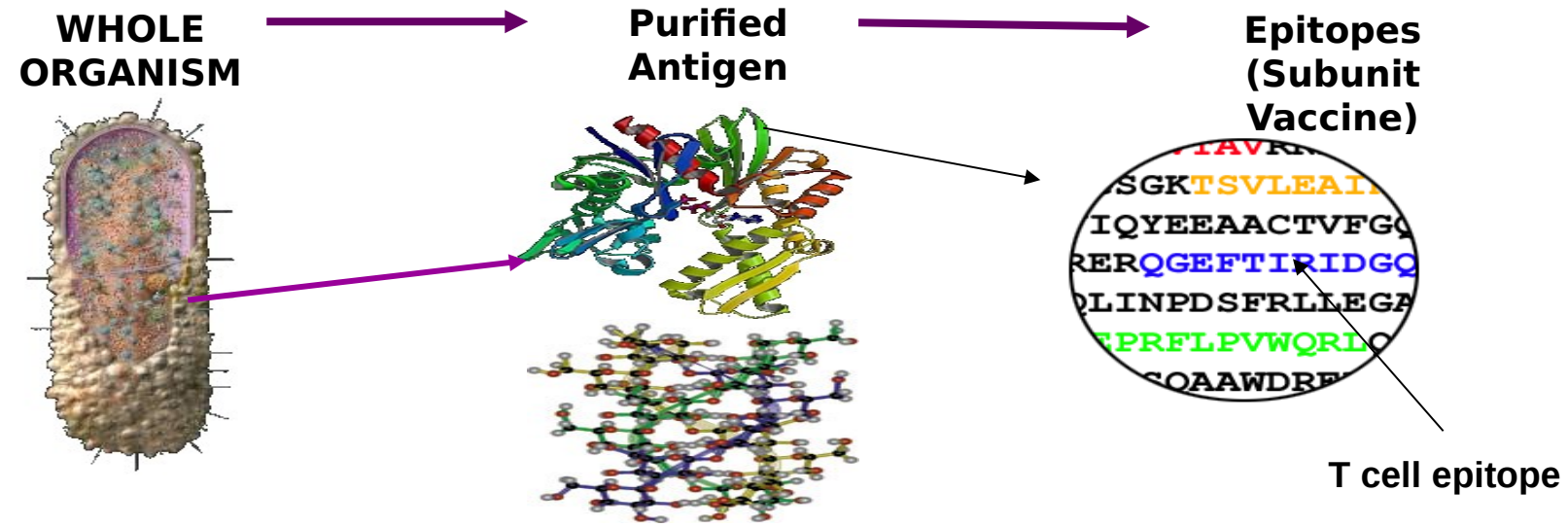


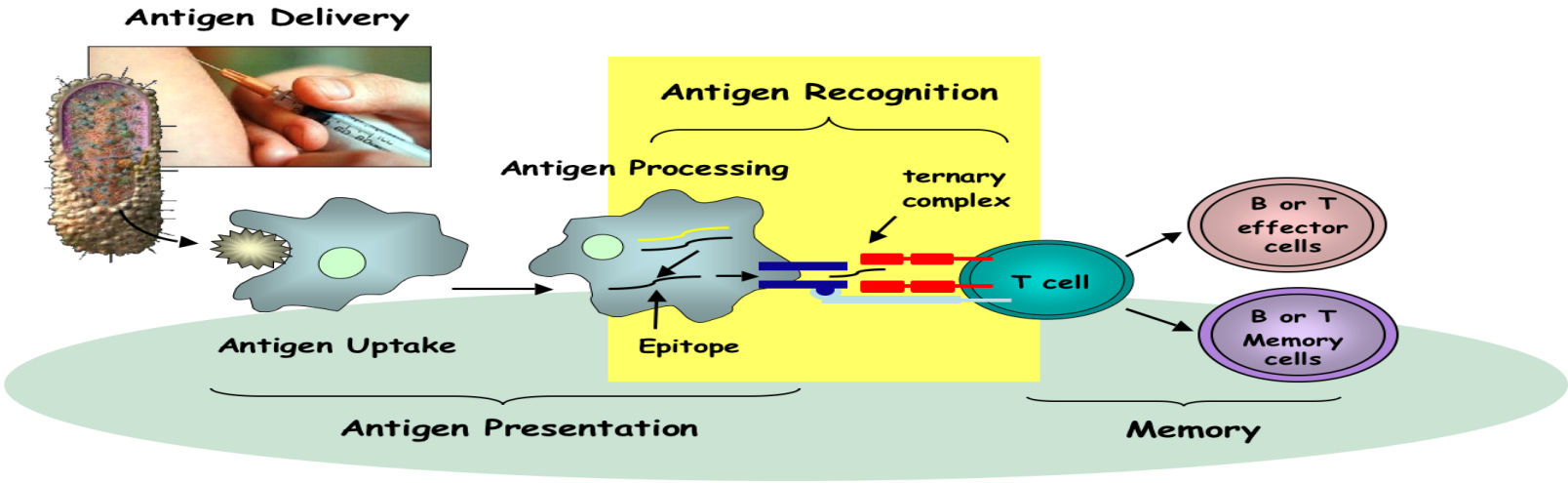
Toxicity of molecules and macromolecules

Tool	Description (Link)
admetSAR 2.0	Tool to predict the chemical ADMET properties (http://lmmd.ecust.edu.cn/admetsar2/)
ADMETopt	Optimization of lead compounds and ADMET screening (http://lmmd.ecust.edu.cn/admetsar2/admetopt/)
ADMETlab	Web-service for systematic ADMET evaluation of chemicals http://admet.scbdd.com/
DrugMint	Computation of drug-like molecules (https://webs.iiitd.edu.in/oscadd/drugmint/)
MetaPred	Prediction of drug metabolizing CYP450 isoforms (https://webs.iiitd.edu.in/raghava/metapred/)
SwissADME	Tool to assess pharmacokinetics, drug-likeness and related parameters of small molecules (http://www.swissadme.ch)
vNN	Webserver for ADMET predictions https://vnnadmet.bhsai.org/
ADVERpred	Web-service for prediction of adverse effects of drugs (http://www.way2drug.com/adverpred/)
BTXpred	Prediction of bacterial toxins (https://webs.iiitd.edu.in/raghava/btxpred/)
ChAlPred	Computation of allergenicity of chemical compounds (https://webs.iiitd.edu.in/raghava/chalpred/)
CTD	The comparative toxicogenomics database (http://ctdbase.org/)
eToxPred	Calculate toxicity and synthetic accessibility of compounds (https://github.com/pulimeng/etoxpred/)
NTXpred	Webserver for predicting neurotoxins (https://webs.iiitd.edu.in/raghava/ntxpred/)
Pred-hERG	Computational tool for predicting cardiac toxicity (http://labmol.farmacia.ufg.br/predherg/)
Pred-Skin	Web portal for accurate prediction of human skin sensitizers (http://labmol.com.br/predskin/)
ProTox-II	Prediction of toxicity of chemicals (http://tox.charite.de/protox_II)
SIDER 4.1	Database on marketed medicines and their adverse drug reactions (http://sideeffects.embl.de/)
ToxiPred	Prediction of aqueous toxicity of small chemical molecules (https://webs.iiitd.edu.in/raghava/toxipred/)
ToxinPred	Estimating toxicity of proteins (https://webs.iiitd.edu.in/raghava/toxinpred/)


Informatics around Healthcare
Bioinformatics
Cheminformatics
Pharmacoinformatics
Clinical informatics
Immunoinformatics
Healthcare
Drug Discovery
• Peptide-based
• Small-chemical based
Toxicity and Adverse Effects
Vaccine Development
• Innate immunity
• Adaptive immunity
Disease Biomarkers
• Predictive
• Prognostic
• Imaging
Healthcare IoTs
• Mobile Apps
• Telemedicine
• Sensor-based
Organization of the Review

Biomolecules Based Vaccines





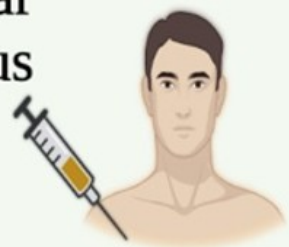
BacVacDB

MECHANISM

- Innate response
- Presentation
- Adaptive response
- Immunity

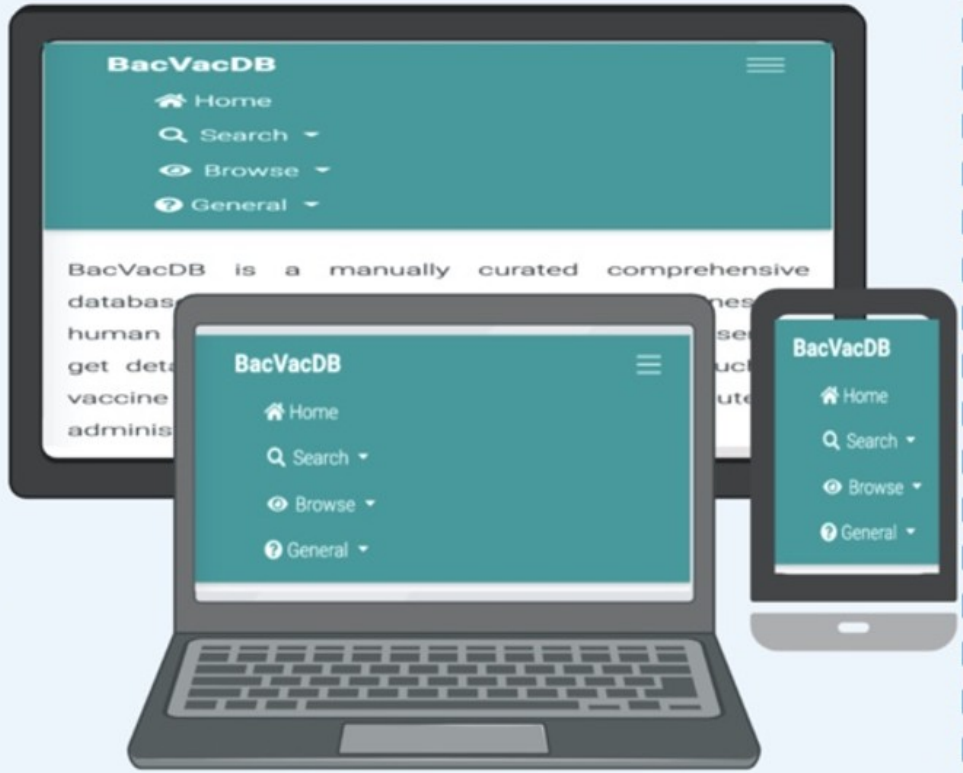
ADMINISTRATION

- Intramuscular
- Subcutaneous
- Intradermal
- Intranasal
- Oral



APPROVAL

- U.S. FDA
- EMA
- CDSCO

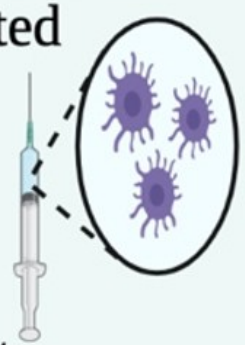


HISTORY

- Chicken cholera (1879)
- BCG (1921)
- Bexsero (2015)

VACCINE TYPES

- Live attenuated
- Toxoid
- Subunit
- Conjugate
- Inactivated
- Recombinant

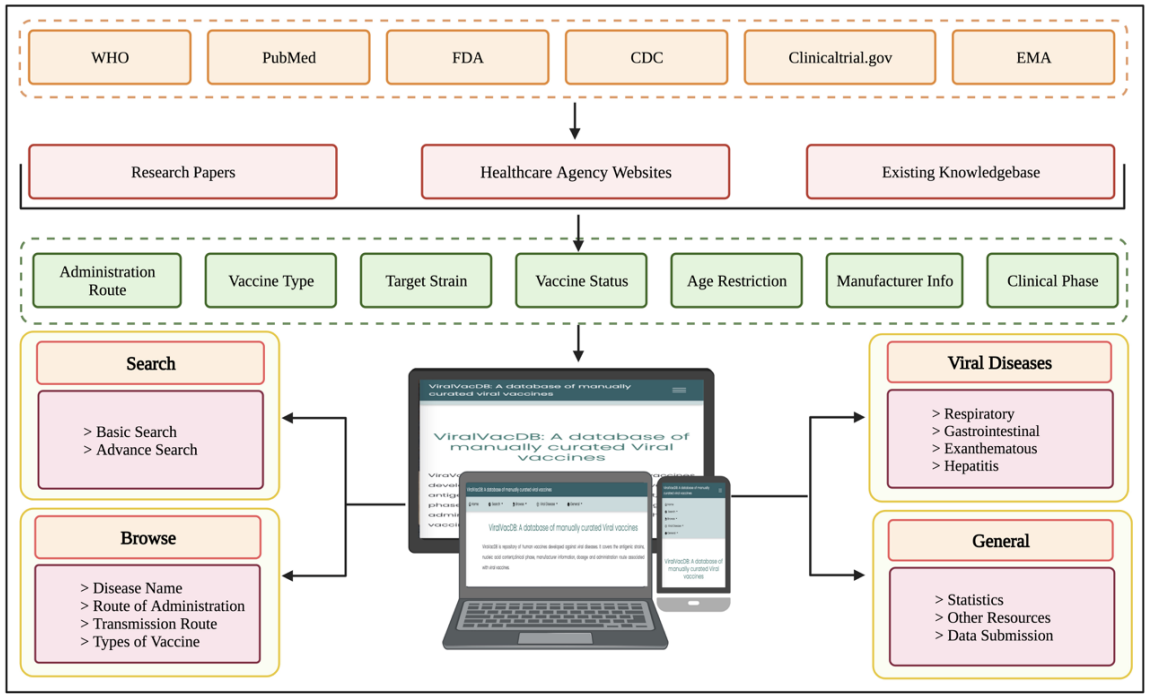


VACCINE STATUS

- Approved
- Clinical trials

ViralVacDB: A database of manually curated viral vaccines

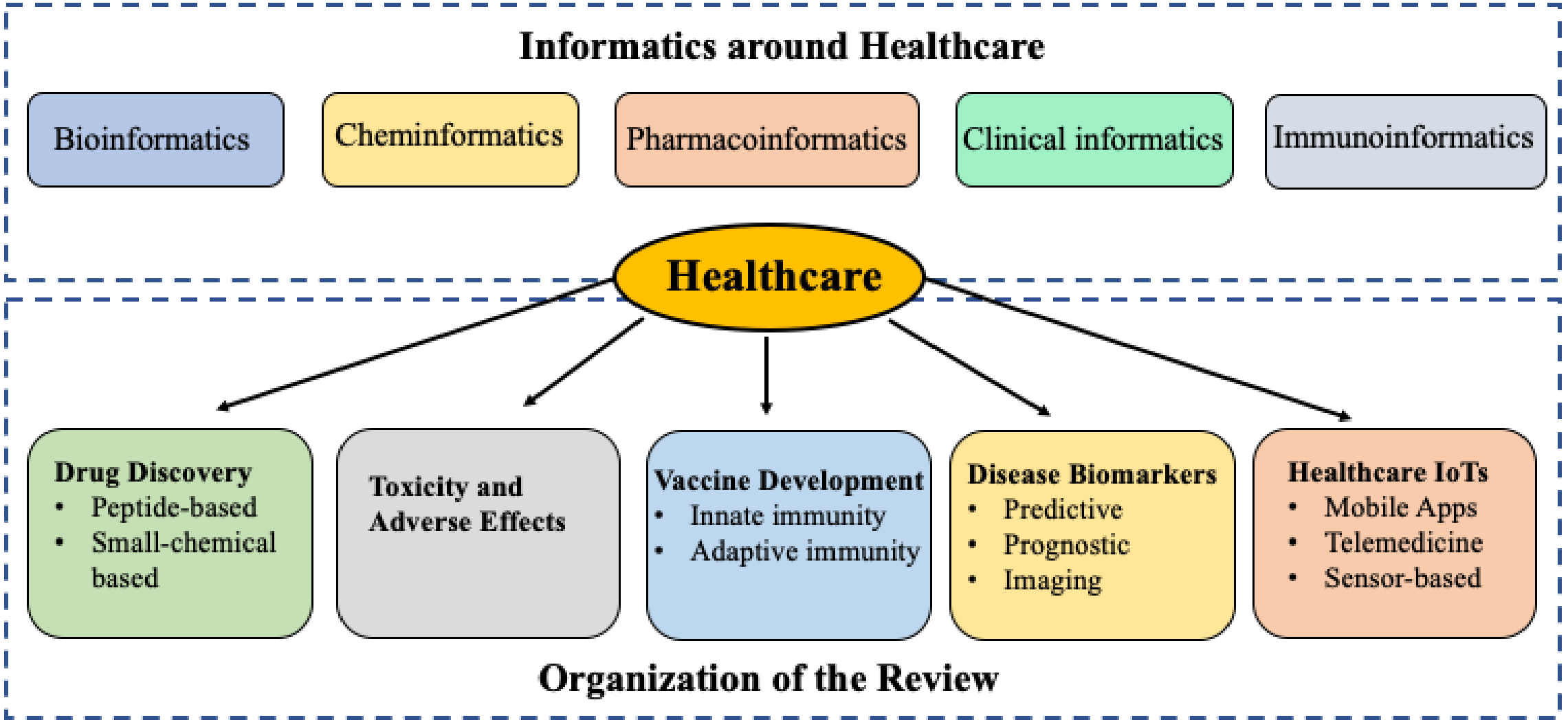
ViraVacDB is a repository of human vaccines developed against viral diseases. It covers the antigenic strains, nucleic acid content, clinical phase, manufacturer information, dosage and administration route associated with viral vaccines.



- **Vaccine against 24 viral diseases**
- **421 viral vaccines**
- **Various routes of administration**
- **Manually curated**

Computer-aided vaccine design

Tool	Description (Link)
pVACtools	Computational toolkit to identify cancer neoantigens (https://pvactools.readthedocs.io/)
IL-6pred	Prediction of interleukin-6 inducing peptides (https://webs.iiitd.edu.in/raghava/il6pred/)
IEDB	Database of immune epitopes and analysis resources (http://www.iedb.org/)
NetMHCIIpan 3.2	Prediction of MHC class II molecules (http://www.cbs.dtu.dk/services/NetMHCIIpan-3.2/)
NetMHCpan 4.0	Prediction MHC class I molecules (http://www.cbs.dtu.dk/services/NetMHCpan-4.0/)
IL-10pred	Identification of interleukin-10 Inducing peptides (https://webs.iiitd.edu.in/raghava/il10pred/)
IL4pred	Prediction of interleukin-4 Inducing peptides (https://webs.iiitd.edu.in/raghava/il4pred/)
IFNepitope	Interferon-gamma inducing epitopes (https://webs.iiitd.edu.in/raghava/ifnepitope/)
LBtope	Predicting linear B-cell epitopes (https://webs.iiitd.edu.in/raghava/lbtope/)
IgPred	B-cell epitopes for different class of antibodies (https://webs.iiitd.edu.in/raghava/igpred/)
DiscoTope 2.0	Discontinuous B cell epitopes (http://www.cbs.dtu.dk/services/DiscoTope/)
Cbtope	Prediction of conformational B-cell epitope (https://webs.iiitd.edu.in/raghava/cbtope/)
ABCpred	ANN based method for predicting B cell epitopes (https://webs.iiitd.edu.in/raghava/abcpred/)
Pcleavage	Proteasomal cleavage sites in a protein (https://webs.iiitd.edu.in/raghava/pcleavage/)
CTLPred	A direct method for prediction of CTL epitopes (https://webs.iiitd.edu.in/raghava/ctlpred/)
TAPPred	Prediction of TAP binding peptides (https://webs.iiitd.edu.in/raghava/tappred/)
ProPred1	Promiscuous MHC Class-I binding sites (https://webs.iiitd.edu.in/raghava/propred1/)
ProPred	Promiscuous MHC Class-II binding sites (https://webs.iiitd.edu.in/raghava/propred/)
PRRpred	Prediction of pattern-recognition receptors (https://webs.iiitd.edu.in/raghava/prrpred/)
PRRDB 2.0	Pattern-recognition receptors and their ligands (https://webs.iiitd.edu.in/raghava/prrdb2 /)
VaxinPAD	Designing peptide-based vaccine adjuvants (https://webs.iiitd.edu.in/raghava/vaxinpad/)
VLCvirus	Lung cancer associated viruses (https://webs.iiitd.edu.in/raghava/vlcvirus/)
CoronaVIR	Web resource for COVID19 (https://webs.iiitd.edu.in/raghava/coronavir/)
Cancertope	Cancer vaccines (https://webs.iiitd.edu.in/raghava/cancertope/)
EbolaVCR	Web resource for Ebola Virus (https://webs.iiitd.edu.in/oscadd/ebola/)


Informatics around Healthcare
Bioinformatics
Cheminformatics
Pharmacoinformatics
Clinical informatics
Immunoinformatics
Healthcare
Drug Discovery
• Peptide-based
• Small-chemical based
Toxicity and Adverse Effects
Vaccine Development
• Innate immunity
• Adaptive immunity
Disease Biomarkers
• Predictive
• Prognostic
• Imaging
Healthcare IoTs
• Mobile Apps
• Telemedicine
• Sensor-based
Organization of the Review

Briefings in Functional Genomics

Volume 20, Issue 4
July 2021

Article Contents

Abstract

Introduction

Resources on cancer genomics

Biomarkers tools for cancer

EDITOR'S CHOICE

Computational resources for identification of cancer biomarkers from omics data FREE

Harpreet Kaur, Rajesh Kumar, Anjali Lathwal, Gajendra P S Raghava ✉ Author Notes

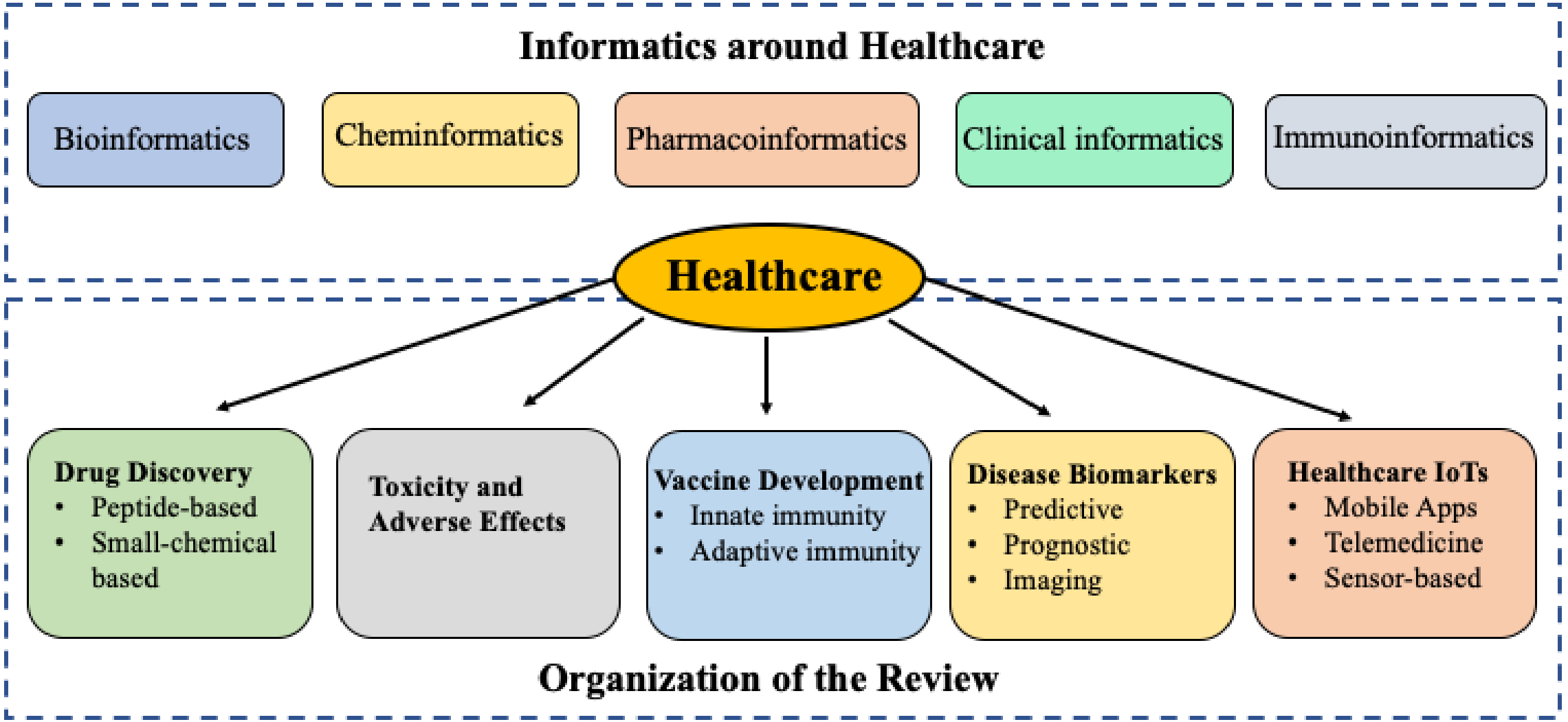
Briefings in Functional Genomics, Volume 20, Issue 4, July 2021, Pages 213–222,
https://doi.org/10.1093/bfgp/elab021
Published: 01 April 2021 **Article history** ▾

Abstract

Cancer is one of the most prevailing, deadly and challenging diseases worldwide. The advancement in technology led to the generation of different types of omics data at each genome level that may potentially improve the current status of cancer patients. These data have tremendous applications in managing cancer effectively with improved outcome in patients. This

Disease biomarkers based on genomic profiles

Tool	Description (Link)
CancerCSP	Classification of early and late stages of CCRC (https://webs.iiitd.edu.in/raghava/cancercsp/)
CancerEnD	Database of cancer associated enhancers (https://webs.iiitd.edu.in/raghava/cancerend/)
CancerLivER	Repository of liver cancer-specific biomarkers (https://webs.iiitd.edu.in/raghava/cancerliver/)
CancerLSP	Identification of early-stage liver cancer patients (http://webs.iiitd.edu.in/raghava/cancerlsp/)
CancerTSP	Prediction of early-stage thyroid cancer (http://webs.iiitd.edu.in/raghava/cancertsp/)
CRC-EBD	Database for epigenetic biomarkers on colorectal cancer (http://www.sysbio.org.cn/EBD/)
MarkerDB	Database of molecular biomarkers (https://markerdb.ca/)
PTSDDB	Database for post-traumatic stress disorder biomarkers (https://ptsd.scai.fraunhofer.de/)
CMcrpred	Survival risk in cutaneous melanoma patients (https://webs.iiitd.edu.in/raghava/cmcrpred/)
CRCRpred	Prediction of risk scores of colorectal cancer patients (https://webs.iiitd.edu.in/raghava/crcrpred/)
SKCMhrp	Classification of high-risk cutaneous melanoma patients (https://webs.iiitd.edu.in/raghava/skcmhrp/)
SurvExpres	Survival analysis from expression data (http://bioinformatica.mty.itesm.mx/SurvExpress)
ePAD	Quantitative imaging biomarkers of cancer treatment response (http://epad.stanford.edu/)
LIFEx	Freeware to extract radiomic features (https://www.lifexsoft.org/)
TCIA	Webserver to create an easily accessible archive of cancer images (https://www.cancerimagingarchive.net/)
3D Slicer	Medical image computing (https://www.slicer.org/)


Informatics around Healthcare
Bioinformatics
Cheminformatics
Pharmacoinformatics
Clinical informatics
Immunoinformatics
Healthcare
Drug Discovery
• Peptide-based
• Small-chemical based
Toxicity and Adverse Effects
Vaccine Development
• Innate immunity
• Adaptive immunity
Disease Biomarkers
• Predictive
• Prognostic
• Imaging
Healthcare IoTs
• Mobile Apps
• Telemedicine
• Sensor-based
Organization of the Review

Mobile apps for Health & Telemedicine

Tool	Description (Link)
1mg	An online pharmacy with a wide range of prescription (https://www.1mg.com/)
Aarogya Setu	Fight against coronavirus infection in India (https://www.aarogyasetu.gov.in/)
AIIMS-WHO CC ENBC	App for nursing colleagues and neonatologist across small hospitals (https://play.google.com/store/apps/details?id=drdeorari.aiims.enc)
Babylon health	All-in-one healthcare on phone (https://www.babylonhealth.com/)
CoDysAn	A telemedicine tool for patients (http://www.codysan.eu/)
Comarch Healthcare	A wide range of healthcare solutions (https://www.comarch.com/healthcare/)
eSanjeevaniOPD	A teleconsultation service ttps://esanjeevaniopd.in/)
Lybrate	Consult doctors online (https://www.lybrate.com/)
ManageMyHealth	Comprehensive telemedicine solution (https://www.managemyhealth.co.nz/m/)
National Health Portal India	A single point access for health information (https://www.india.gov.in)

Thank