**Simon Kasif Key Work/Research Contributions Timeline**

**(both solo and group)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | Work | General AI/Biology/Medicine Topic | With | Color scheme of Innovation)  AI (green)  CS (red) Innovation  Biology (blue) | |
|  | \*Green: An early or pioneering contribution  \*Considered a Textbook Result  \*Transformative Community Work or Widely used Systems  **Bold – highly cited or visible** |  | **Red: Lab/Group Member**  **\***Blue: Collaborator |  | |
| 1981 | Pyramid Linking Converges | Early Unsupervised Deep Learning (AI theory) | A. Rosenfeld | XX | |
| 1983 | First implemented Parallel AI Inference (Datalog) System on a distributed memory machine  The system was using fork – join, an early and rudimentary implementation of Map – Reduce for parallel AI | Parallel Joins Coupled with Datalog (AI system) | J. Minker et al | XX | |
| 1983 | First Parallel Logic Programming AI System to be mapped and tested on a multiprocessor (ZMOB) | Parallel AI (system) | J. Minker et al | XX | |
| 1985 | Founder, Johns Hopkins AI Laboratory | Initially Parallel AI, Scalability, Logical Inference, Networks)  Later Machine Learning and Computational Biology | S. Salzberg | X X X | |
| 1985 | Arc Consistency / Discrete Relaxation is Inherently Sequential  **FOUNDATIONAL THEORY AND HIGHLY CITED** | Parallel AI  (AI theory) | Sole author | XX | |
| 1985-1987 | Graph Separator based Parallel SAT Solving | Parallel AI  (AI theory) | J. Reif  D. Sherlekar | XX | |
| 1987 | Provably Optimal Parallel Term Matching and Anti-Unification | Parallel AI  (AI theory) | Art Delcher | XX | |
| 1991 | An early introduction of Data Streaming into ML and Data Mining  Mostly theory. Upper, Lower Bounds for Data Streaming  Application in Cognitive Science | Preceded by several years the early Data Streaming Results that popularized the area (AI Theory) | D. Heath,  R. Kosaraju,  S. Salzberg | XX | |
| 1990 | **223 moves to capture a piece and win**  **\*Considered Texbook and seminal work in computer chess**  **Follow up on Ken Thompson’s 51 moves to win** | Parallel AI System on the Connection Machine Revolutionized chess  (AI system) | L. Stiller PhD Thesis | XX | |
| 1993 | Introducing Randomization in Decision Tree (DT) Induction Prior to Random Forests  **HIGHLY CITED** | Randomization in DTs, Random Projections (system) | D. Heath,  S. Murthy,  R. Beigel,  S. Salzberg | XX | |
| 1993- | **Bayes Networks and Biology**  **HIGHLY CITED** | Early proposal for causal BNs in Biology | A. Delcher et al | X X X | |
| 1993- | **Rudimentary generative AI in biology: a small but early step towards using AI in synthetic biology** | First of a kind Application of BN for Biology: | A.Delcher et al | X X X | |
| 1994 | **OC1 – Widely Used Open Access Decision Tree System**  **HIGHLY CITED** | (Randomization, Scalability by Sorting on Attributes, Ensemble of DTs,  AI system) | S. Murthy,  S. Salzberg | XX | |
| 1994 | Logarithmic Time Queries and Updates in Probabilistic Networks  Factorization Extending on Miller-Reif in Bayes Nets | Novel factorization and compilation of BNs enabling parallelism and dynamic updates  Application to In-silico mutagenesis  First application of graphical models to synthetic biology | J. Pearl  A. Grove, A. A. Delcher | X X X | |
| 1994-1996 | First of a kind Bayes Nets Kernels  (very early and perhaps the first of a kind EMPIRICAL work but lacks mathematical rigor found in the follow-up papers) | Producing a BN Kernel, integrating BN-s with k-NN, Testing on ML Benchmarks  (system and theory) | J. Rachlin,  D. Waltz  S. Salzberg | X X X | |
| 1996 | **AAAI Symposium Learning Complex Behaviors** | Moving Machine Learning from Toy Problems to Learning Complex Behaviors and Systems | S. Russell  M. Jordan  D. Koller  Over 100 scientists | XX | |
| 1997 | **Human Centered Systems: Information, Interactivity and Intelligence** | Launched the Human Centered Intelligence Initiative at NSF  Early Warning about AI Safety and the problem of Control | J. Flannagan  T. Huang  P. Jones  over 100 scientists | XX | |
| 1998 | Computational Methods in Molecular Biology”, Elsevier Publ.  Focus on AI and Biology including Bayes Nets  Some textbook material – not all. | One of the earliest books in Computational Molecular Biology | S. Salzberg  D. Searls | X X X | |
| 1998 | **GLIMMER: widely used Open Access AI System for Microbial Annotation**  **First variable length models in Bio**  **HIGHLY CITED** | Millions of new enzymes identified with massive impact on science and biotechnology  (open access system) | A. Delcher  S. Salzberg O. White | X X X | |
| 1999 | **MUMMER: widely used and early system for whole genome comparison (bacteria)**  **First open access whole genome bacterial analysis**  **HIGHLY CITED** | Wide use in clinical and biological research with exceptional impact on biology and disease  (open access system) | A.Delcher  S. Salzberg | X X | |
| 1999 | Multiplex PCR for Gap Closing in Genomes | Bridge between Complex Combinatorics and Genomics  (theory and experimental implementation) | S. Salzberg  H. Tettelin | X X | |
| 1999-2001 | **Human Genome Project**  **HIGHLY CITED** | **COMPUTATIONAL ANALYSIS TEAM** | Lander et al | XX X | |
| 2000 | **DARPA SYSTEMS BIO AND SYNT. BIO SYMPOSIUM** | LANGUAGE MODELING FOR BIOLOGY | Co-organizer with Bob Berwick et al  Over 100 scientists attended | X X X | |
| 2000 | Bayes Nets and Genomic Integration |  | V. Pavlovic | X X X |
| 2002 | Center for Advanced Genomic Technology (at BU) |  | Charles Delisi |  |
| 2003-2005 | **Network based Gene Function Prediction**  **Wide impact on most popular GENE FUNCTION PREDICTION systems today HIGHLY CITED** | Introducing Network Propagation into Function Prediction  Established DISCIPLINE | S. Letovsky | X X X | |
| 2004 | **Hopfield Networks and Gene Function Prediction** | Introducing Network propagation into Function Prediction | CM. Ding  T.M. Murali  C. Cantor | XXX | |
| 2004-2020 | **Direct Involvement in Experimental Validation of Computational Predictions**  **(among the early pioneers of computational / experimental hybrid projects)**  **HIGHLY CITED** |  | R. Roberts  C. Cantor  G. Cooper  M. Steffen  R. Jain  B.Bernstein  A. Regev  R. Kahn  S. Tseng  ME. Patti  Combrex  T. Gardner  J. Collins | X | |
| 2004 | Learning Hidden Matchings in Graphs from queries | New mini-area in graph theory | N. Alon  R. Beigel et al | XXX | |
| 2004 | **Starting COMBREX CONSORTIUM**  **Academy Meeting**  **(AI Driven Community Science)** | A Community Project using Active Learning for Recommending and Driving Experiments | Rich Roberts & > 50 scientists | XXX | |
| 2005 | Multi-nodes graphs and Multiplex PCR  **MUPLEX System** | Multiplex PCR analysis and open access system  Motivated by non-invasive detection of fetal DNA in maternal blood (later liquid biopsy) | J. Rachlin  CM. Ding  C. Cantor  N. Alon | XXX | |
| 2006 | Biological Context Networks (BCN)-s | Pioneered a mini new area of network science | V. Asodi  J. Rachlin  N. Alon | XXX | |
| 2007- | Network Biology of Wellness  (many talks given) | In progress | R. Kahn  I. Kohane | XXX | |
| 2007 | **Inference of Regulatory Networks in Bacteria with Validation**  **HIGHLY CITED** | >1700 citations | J. Faith  T. Gardner  J. Collins | XXX | |
| 2000 - | **Popularizing Data Integration in Biological Science Using Networks** | Large field now | With many independent contributions | XXX | |
| 2010-2014 | Using Machine Learning (AI) for Drug Monitoring in Social Medial (Twitter) |  | Clark Freifeld  J. Brownstein at al |  | |
| 2003- | Active Learning for Biology  THE COMBREX PROJECT |  | Rich Roberts et al | XXX | |
| 2010-2017 | Director, Regional Bioinformatics and Systems Biology Core (the core contributed to many papers in Nature Medicine, Nature Cell, Cell Metabolism papers in major journals) | NIH Center |  | XXX | |
| 2012 | Driving Citizen Science with AI | NIH Common Fund Meeting | L. Griffith Meyer  S. Seung | XXX | |
| 2013 | **The COMBREX REPORT**  **First AI driven community science in Biology (by active learning) paper (2AI2BIO)\*** |  | COMBREX Consortium | XXX | |
| 2014 | **Reprogramming Stem Cells driven by computational analysis and epigenomics**  **FOUNDATIONAL AND HIGHLY CITED** | Highly cited | E. Rheinbay  M. Suva  B. Bernstein  et al | X | |
| 2015 | The Biomed Collaboration Initiative | Harvard Medical School |  | X X | |
| 2010-2020 | The Provenance Initiative | In progress | Rich Roberts et al | X X X | |
| 2020- | AI and Biology Initiative | In progress | Rich Roberts et al | XXX | |
| 2020- | Automating COVID Drug Discovery (small) | In progress |  | XXX | |
| 2020- | Use of Complexity Theory in Pooling COVID-19 TESTS (small) | Theory and system | R. Beigel | XX | |
| 2018- | AI based Genomic Inspectors | System and Theory | Brian Haas  Aviv Regev  et al | XXX | |
| 1996-2020 | Many advisory boards and steering/advisory committees at both local, single university or national or international level including Hopkins Mind/Brain Institute, BU, Harvard Med, CMU, Columbia, Alberta Innovation Academy, U. Chicago, NSF, Joslin Diabetes Center, start-ups, conferences, NIH, NSF, DARPA, industry and more) |  |  |  | |
| 1985-2022 | Over 60 students/fellows mentored or co-mentored, many minorities, women  <http://sites.bu.edu/phenogeno/people/> |  |  |  | |
| 1985-2022 | Over 70 direct small group collaborators (in addition to consortium memberships and large papers)  <http://sites.bu.edu/phenogeno/people/> |  |  |  | |