

RAJDEEP SINGH

# What is Bioinformatics

Copyright © 2021 by Rajdeep singh

*All rights reserved. No part of this publication may be reproduced, stored or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise without written permission from the publisher. It is illegal to copy this book, post it to a website, or distribute it by any other means without permission.*

*First edition*

*This book was professionally typeset on Reedsy.*

*Find out more at [reedsy.com](https://reedsy.com)*

# Contents

1	Abstract	1
2	Keywords	2
3	Introduction	3
4	Definitions	4
5	Multidisciplinary approach	5
6	History	6
7	Scope of bioinformatics	8
8	Where Bioinformatics help	10
9	Software and tools in bioinformatics	11
10	Language of bioinformatics	12
	R Language	12
	Python Language	13
	Other Language	13
	Role of Coding in Bioinformatics	13
	Lab in Bioinformatics	14
	Dry Lab:	14
	Wet Lab	14
11	Operating systems	15
	Why Use of Linux?	16
	Where use of Linux?	16
	Why do we not use Windows or macOS in Bioinformatics?	16
12	Application of Bioinformatics	17
13	Conclusion	18
14	References	19



## Abstract

Bioinformatics is a new branch of the science world. Bioinformatics is a multidisciplinary approach. We use bioinformatics to understand biology information and save it into the biological database.

Apply data science on biological databases, discover a new drug, and modify extinction drugs to improve human life.

## Keywords

Biology, Bioinformatics, Computer science, Software and tools, Coding.

## Introduction

Bioinformatics is a multidisciplinary approach. Bioinformatics works jointly with different science branches to improve understanding and concept of biology.

Bioinformatics has recently been added to the science world. Bioinformatics comes with many opportunities for research to submit new research and build new products to improve human welfare.

The term Bioinformatics was invented by Paulien Hogeweg and Ben Hesper in 1970.

**Margaret Oakley Dayhoff is the mother and father of Bioinformatics.**

Bioinformatics name joins two-part of the word.

Bio means biology.

Informatique (a french word) means **data processing**.

In the Local internet language, if you like biology and coding. that is bio-informatics.

## Definitions

In Bioinformatics, many companies or people give different definitions. In the article, we list out some important definitions.

1. Bioinformatics is when biology, computer science, and information technology merge into a single discipline for managing and analyzing biological data using advanced computing techniques.
2. The combination ( or marriage) of biology and information technology.
3. The stored data is available in the form of sequences and structures of proteins and nucleic acid.
4. The science of collecting and analyzing complex biological data such as genetic codes.
5. It involves the computational tool and method used to manage, analyze and manipulate volumes and volumes of biological data.



## Multidisciplinary approach

The multidisciplinary approach means bioinformatics work with lots of fields like computer science, biology, mathematics, biotechnology, statistics, biochemistry, etc.

## History

Over a century ago, bioinformatics history started with an Austrian monk named Gregor Mendel. He is known as the “Father of Genetics.” He clearly defined statistics and data studies on plants. After Mendel, many new scientists came and tried to find discoveries in the human world.

In the article, We discuss some major discoveries in the world.

- The term bioinformatics was first introduced by Paulien Hogeweg and Ben Hesper in 1970.
- In 1972, Paul berg made the first recombinant DNA molecule using the ligase. That same year, Stanley Cohen, Annie Chang, and Herbert Boyer produced the first recombinant DNA organism.
- Joseph Sambrook led a team that refined DNA electrophoresis using agarose gel in 1975.
- By 1977, a method for sequencing DNA was discovered, and the first genetic engineering company, Genetech, was founded.
- In 1978, David Botstein discovered Restriction fragment

length polymorphisms ( RFLP).

- In 1980, Kary Mullis invented polymerase chain reaction (PCR)
- In 1982, the Genetics computer group (GCG) was created as part of the University of Wisconsin biotechnology center.
- In 1986, The SWISS-PORT Database was created by the Department of medical biochemistry of Geneva and the European molecular biology laboratory ( EMBL).
- In 1987, the use of yeast artificial chromosomes (YAC) Was described.
- In 1987, the human genome initiative was started.
- In 1990, the BLAST program was implemented.
- In 1991, the first time the term bioinformatic appeared in the scientific literature.
- In 1992, a genetics linkage map of the entire human genome was published.
- In 1996, Affymetrix produced the first commercial DNA chips.
- In 1996, the cloned sheep Dolly was born on July 5, 1996.
- In 1998, the Swiss Institute of Biotechnology was established as a non-profit foundation.
- In 1999, the human chromosome was completely sequenced.
- In 2000, international research was sequenced.
- In 2000, the international research consortium published chromosome 21.
- In 2001, the human genome ( 3000 Mbp) was published
- In 2003 the human genome project was completed in April 2003
- In 2005, the genome sequence of rice was complete

## Scope of bioinformatics

Bioinformatics is a multidisciplinary approach. Bioinformatics combines many scientific fields. For that reason, the bioinformatics scope is extensive. Bioinformatics cover all major fields.

- Molecular Medicine
- Drug Therapy
- Microbial genome applications
- Forensic analysis of microbes
- Bio-weapon creation
- Biotechnology
- Biological computation
- Bio-engineering
- Biology
- Artificial intelligence
- Data mining
- Image processing
- Computational biology
- Statistics
- Mathematics

## SCOPE OF BIOINFORMATICS

- Molecular biology
- Genetics

## 8

### Where Bioinformatics help

- Experiment molecular biology
- In genetics and genomics
- Generating biological data
- Analysis of gene and proteins expression
- Comparison genomic data
- Simulation & Modeling of DNA, RNA & Protein

## Software and tools in bioinformatics

In bioinformatics lots of tools in the market, some free (open-source) and paid. In this field, I watch very few software as paid rest of free and open-source.

- BLAST
- FAST
- BioJS
- Biopython
- Bali-Phy
- BioPHP

### **All tools or software**

[https://en.wikipedia.org/wiki/List\\_of\\_bioinformatics\\_software](https://en.wikipedia.org/wiki/List_of_bioinformatics_software)

### **Open-Source software**

[https://en.wikipedia.org/wiki/List\\_of\\_open-source\\_bioinformatics\\_software](https://en.wikipedia.org/wiki/List_of_open-source_bioinformatics_software)

## Language of bioinformatics

In Bioinformatics 2, language is most commonly used it. But according to the company Requirements. Your programming language also changes. this recommendation is my expression base

1. R Language
2. Python Language
3. Other Language

### R Language

**R** is a programming language. R is also used for statistical computing and graphics in build support by the R Foundation. The R language is widely used among statisticians and data miners for developing statistical software.



## Python Language

Python is a general-purpose, high-level language. Python is used everywhere in the world. You use python and build anything. Python is used for machine learning, ai, data science, web development, desktop app design. Currently, python also rapidly grows in statistical analysis.

## Other Language

In another language category, we use the bash, C, and c++ language most commonly. But most of the time, real-world use of the R language of bio-information may change company to company based on company requirements. But Now, Days, python also grows rapidly.

## Role of Coding in Bioinformatics

Two types of bioinformaticians in this world. First, which person knows about coding or software building skill, and second is the only use of software to predict and maybe show some kind result based on data use of the software. A non-coding person does not build their tool or software. They use other people to build software or tool.

## Lab in Bioinformatics

Common two types of lab consider in bioinformatics science.

Dry lab

Wet lab

### Dry Lab:

In Dry lab, we use computers to perform prediction based on biological data. We don't use chemicals, instruments and do not extract any DNA from cells. We use a laptop or desktop and good qualitative internet to analyze the data and give a prediction model.

### Wet Lab

We use chemicals, cells, and instruments to get information and store it in the database inside the wet lab. With the help of these data, bioinformatician builds and predict the model base on data.

## Operating systems

The operating system is a basic fundamental tool for computer or laptop contact with the hardware. The most common operating system is three types.

1. Window
2. Mac OS
3. Linux Kernel

Window and Mac Os use the daily base widely used operating system. Window and Mac Os also provide a good UI interface for humans. Another hand Linux Kernel does not provide any UI interface. But Linux provides a command-line tool, which is more than important for us.

Linux also supports many distributions. Linux distributions help to provide a good UI interface for us. The most Commonly Linux distribution is Ubuntu.

## Why Use of Linux?

Linux is very small in size and also runs with a small size of Ram and Storage.

## Where use of Linux?

Right Now Linux Every Where. You use Linux to create a project with low device storage and Ram. Linux is a free and open-source project.

Most people Linux use in Android Mobile. Android is also built with Linux.

Linux is used to deploy websites on the Cloud. Every Cloud provides support for all Linux distributions.

## Why do we not use Windows or macOS in Bioinformatics?

It depends from person to person and company to company. Most companies and persons like Linux in daily life, and They build bioinformatics and software also. Another person, like macOS or Windows to builds bioinformatics products. Both of the three systems work the same and design, build bioinformatics softwares.

The downfall of Windows and macOS.

- Large Size
- More Ram And Storge to perform the task.

## Application of Bioinformatics

Bioinformatics workarounds biology and computer science. That reason the bioinformatics application list is huge. But I tell you a few of that, which are very important in bioinformatics

- Sequence mapping of biomolecules ( DNA, RNA, and Proteins).
- Finding of sites that restriction enzymes can cut
- Prediction of functional gene products
- Design 3d Structure of the molecule
- Designing software
- Developing device drivers
- Embedded solution
- Biological data management
- Data and image analysis
- Drug discovery
- Genomics and proteomics

## Conclusion

Bioinformatics is a new branch of the science world. Many opportunities and discoveries in this field are pending.

Suppose you think and are inserted in bioinformatics. That is the right time to start learning. Right now, lots of YouTubers and courses are available on the internet. You just set your mind start learning. Your future is very bright in bioinformatics.

## References

- <https://en.wikipedia.org/wiki/Bioinformatics#History>
- <https://www.biologydiscussion.com/biodiversity/bioinformatics/bioinformatics-history-coverage-components-and-applications/11294>
- <https://academic.oup.com/bib/article/20/6/1981/5066445>
- <https://www.slideshare.net/AsadAfridi5/introduction-to-bioinformatics-147580009>
- <https://www.slideshare.net/SomduttSharma4/bioinformatics-75142518>
- <https://www.slideshare.net/JTADrexel/bioinformatics-2512758>
- <https://www.slideshare.net/vidhyakalaivani29/major-data-bases-in-bioinformatics-71778405>

