

# Exploring Scala/Bio-Scala/ScalaNLP for Advanced Bio-informatics Algorithms R&D w.r.t COVID-19.

Nirmal Tej Kumar

Independent Consultant – Informatics/Photonics/Nanotechnology/HPC R&D.

R&D Collaborator – USA/UK/Israel/BRICS Group of Nations.

Current Member – ante Inst UTD Dallas TX USA.

Contact\_info – [hmfg2014@gmail.com](mailto:hmfg2014@gmail.com)

## **Abstract :**

Scala/Bio-Scala +AI Based Probing of Advanced Bio-informatics → A Simple Suggestion + Short Technical Communication on JVM Environments. Next generation sequencing and vaccine design needs improvement/s. We need to do a lot of R&D w.r.t COVID-19.

**index words :** Scala/Java/JVM/AI/ML/BIG DATA/IoT/HPC – High Performance Computing.

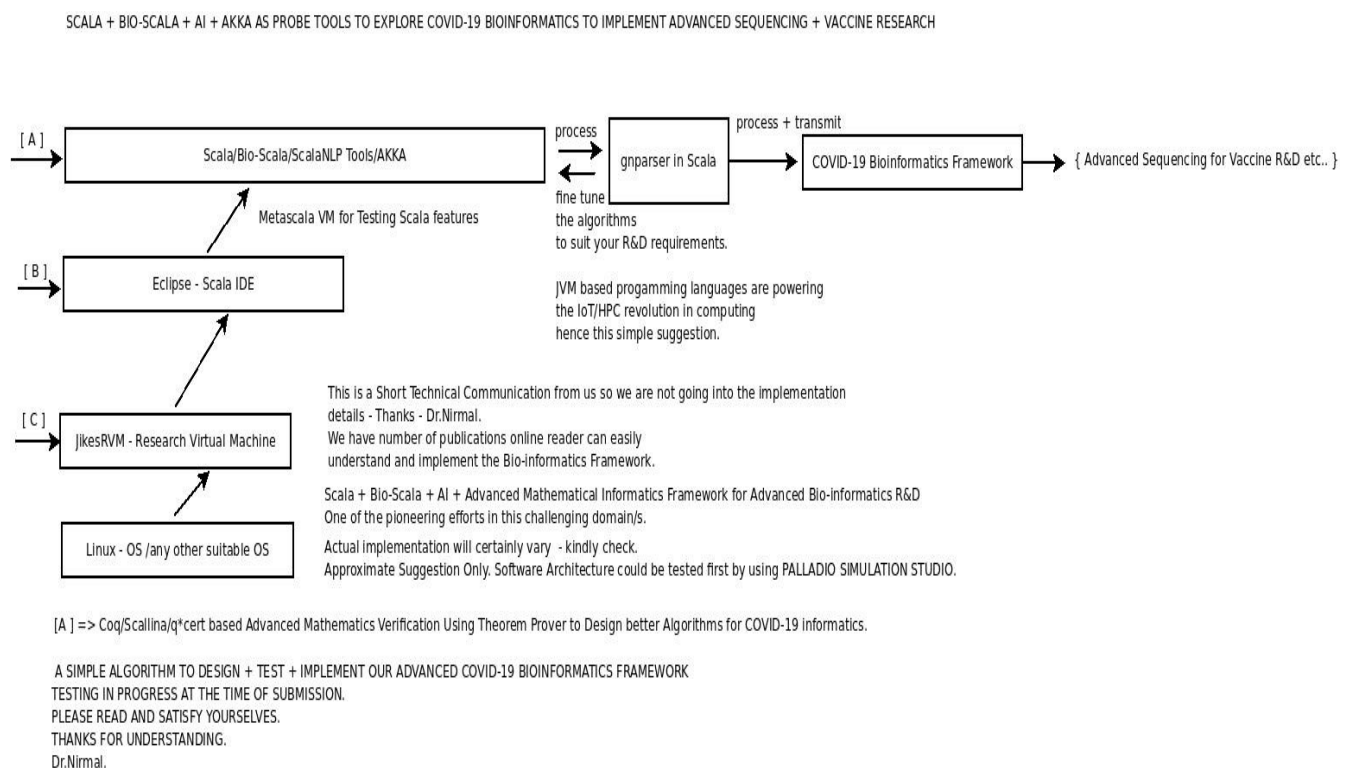
## [I] Main Idea + Inspiration :

“It is a functional and hybrid object-oriented programming language, with emphasis on functional programming preferences like laziness and immutability. Scala evolved to the general public in 2004. Created by Martin Odersky, Scala programming language was released on the Java platform for general-purpose. The Scala is not any whimsical notion, it’s an abbreviation of “**Scalable**”: a language that is scalable as per the needs of users.

The early release of Scala surely gave it the advantage of maturing more over the time and taking the spotlight from the Kotlin programming language.”

[ Source - <https://towardsdatascience.com/kotlin-vs-scala-which-is-right-jvm-for-2020-952bc5b955cc> ]

## [II] Scala/Bio-Scala based R&D Bio-informatics Platform :



[ Figure I -> A Simple Algorithm -> Scala in Bio-informatics R&D Framework ]

### **[III] R&D Discussion + Simple Analysis :**

“ The goal of the Bio-Scala project is to create a scalable and functional bio-informatics library; which can be used from Scala, Java, JRuby, Jython and Clojure; building on existing libraries available on the JAVA Virtual Machine (JVM), including Bio-JAVA and Bio-Ruby. This is the place for the Bio-Scala project under the liberal BSD license.

This edition of Bio-Scala has:

- Strongly typed DNA, RNA Sequence, gapped and IUPAC ambiguous Sequence
- Sequences contain lists of Nucleotide, AminoAcid, IUPAC, Gapped, Codon
- Flexible Sequence user defined attributes
- Transcription of DNA to RNA
- Translation of (ambiguous) RNA to AA (using BioJava3)
- Formats:
  - . Iterator based FASTA reader and writer
  - . Iterator based PAML (PHY) file format reader.”

[ Source - <https://github.com/bioscala/bioscala> ]

**We intend to fine tune gnparser to probe our Bio-informatics needs in the Context of COVID-19 Using Smart Devices + IoT + HPC R&D Environment/s.**

“gnparser”: a powerful parser for scientific names based on Parsing Expression Grammar “

[ Source - <https://pubmed.ncbi.nlm.nih.gov/28549446/> ]

[ Source - 2017 May 26;18(1):279. doi: 10.1186/s12859-017-1663-3 ]

### **[IV] Acknowledgment/s :**

Sincere Thanks to all WHO made this happen in my LIFE. Non-Profit R&D.  
Inspiring Others Always is GOOD.

### **[V] Conclusion/s With Future Perspectives :**

Scala is the NEXT BIG THING in Advancing IoT /HPC Revolution Using BIG DATA + AI + NLP + Parsers for example “gnparser” written in Scala to Probe the Frontiers of Advanced Bio-informatics. Hence, this Short Technical Communication highlighting the important aspects of JVM + JVM based Programming Languages. To the best of our knowledge, this is one of the pioneering R&D efforts in this challenging domain worth exploring.

## [VI] Important References :

- [a] <http://scala-ide.org/>
- [b] <https://arxiv.org/pdf/1807.06808.pdf>
  
- [c] <https://github.com/tejdnk-2019-ShortNotes/tejdnk-Space-Medicine-Informatics-github.io>
- [d] <https://scala-lms.github.io/>
  
- [e] <https://hedleyproctor.com/2014/11/writing-a-dsl-with-scala-parser-combinators/>
- [f] <https://github.com/sterglee/scalalab>
  
- [g] <https://github.com/bioscala/bioscala>
- [h] <https://github.com/JBakouny/Scallina>
  
- [i] <https://coq.inria.fr/refman/practical-tools/coqide.html>
- [j] <https://querycert.github.io/>
  
- [k] <http://hirzels.com/martin/papers/sigmod17-qcert.pdf>
- [l] <http://lambda-the-ultimate.org/node/4794> && <https://github.com/lihaoyi/Metascala>
  
- [m] <https://towardsdatascience.com/introduction-to-spark-nlp-foundations-and-basic-components-part-i-c83b7629ed59>
  
- [n] <https://www.quora.com/Why-should-I-learn-Scala-for-machine-learning>
  
- [o] <https://akka.io/> && <http://www.scalanlp.org/> → ScalaNLP Information + Notes.
  
- [p] <https://medium.com/activewizards-machine-learning-company/top-15-scala-libraries-for-data-science-in-2018-4b2cb5c5367e>

[ THE END ]