Technological advancement enables low cost and high-throughput sequencing (HTS) data, such that, once the study deviates slightly from a standard work flows, custom scripts are in demand for faster data processing and better visualization. The developed Python library is called HTSeq, which includes parsers for common file formats and thus making it suitable to work as general platform for writing custom scripts for analyzing HTS.

HTSeq provides parsers for sequences, short reads, short-read alignments, genomic features, annotation, and score data. Each parser is provided as a class, such that those objects are tied to a file and work as iterator generator. This architecture allows the parser to be used in the head of a *for* loop and generates a sequence of record objects that are analyzed by the loop variables. By recording sequences outside the loop, HTSeq significantly reduces time taken to analyze HTS as well as the computational cost. In order to further enhance performance and compatibility, majority of HTSeq is written in Cython, a tool that translates Python code augmented with type information to C, to reduce dependencies.

With this being said, Author developed HTSeq with an aim to benefit for researchers who needs develop codes entirely from scratch to process certain HTS data.