**The definitive Explanation on how Spring Batch could may explain How Crispr cas9 Works**

**Author : Wadï Mami**

**Date : 16/01/2023**

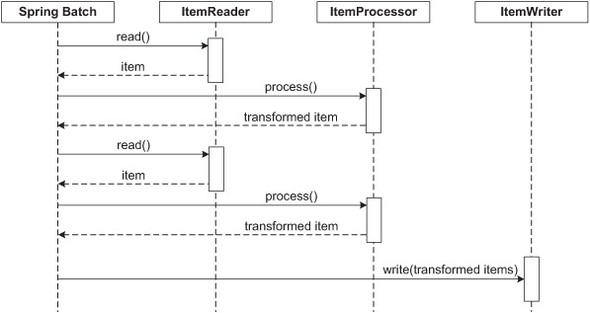
**Email :** [**wmami@steg.com.tn**](mailto:wmami@steg.com.tn) **/** [**didipostman77@gmail.com**](mailto:didipostman77@gmail.com)

**Abstract :**

Many people were not convinced about how Spring Batch could may explain how CRISPR Cas9 Works. In this following paper I will try to represent the Howto in more details with some changes to fit how CRISPR Cas9 Works as We will have 2 batchs one batch job for creating the CRISPR Arrays with input virus DNA data/DNA.csv and output the CRISPR Arrays in outputFile.csv . And the second Batch is the scond phase disabling the virus by using input the CRISPR Arrays outputFile.csv and altering the virus DNA data/DNA.csv

**The CRISPR arrays :**

**Spring Batch is the bacteria**

.

The bacteria capture snippets of DNA from invading viruses and use them to create DNA segments known as CRISPR arrays

 example private String dna\_pattern = "AATTCC"; //snippets of DNA from invading viruses in

<https://github.com/didipostman/CrisprCas9/blob/main/src/main/java/com/juxtapose/example/ch02/DNA_SequenceProcessor.java>

<=> Spring Batch read DNA file or DNA database , The DNA file or the DNA database are Viruses DNA.

 SpringBatch read() --->ItemReader and ItemReader return item. <=> Spring Batch process() ----> ItemProcessor and

 return transformed item = DNA segments known as CRISPR arrays Here I use DNA\_sequenceProcessor class that

 implements ItemProcessor and uses Karp Rabin (you can use other DNA pattern recognition algorithm)

Below is a code representation of the same concepts shown above:

List items = **new** Arraylist();

**for**(**int** i = 0; i < commitInterval; i++){

Object item = itemReader.read()

Object processedItem = itemProcessor.process(item);

items.add(processedItem);

}

itemWriter.write(items);

The following code fragment shows how to define a CRISPR ARRAYS step in XML:

<job id="dnaSeq">

<step id="CRISPR\_ARRAYS">

<tasklet transaction-manager="transactionManager">

<chunk reader="csvItemReader" writer="csvItemWriter"

processor="DNA\_SequenceProcessor" commit-interval="2">

</chunk>

</tasklet>

</step>

</job>

<bean:bean id="csvItemReader"

class="org.springframework.batch.item.file.FlatFileItemReader"

scope="step">

<bean:property name="resource"

value="classpath:ch02/data/DNA.csv"/>

<bean:property name="lineMapper">

<bean:bean

class="org.springframework.batch.item.file.mapping.DefaultLineMapper">

<bean:property name="lineTokenizer" ref="lineTokenizer"/>

<bean:property name="fieldSetMapper">

<bean:bean class="org.springframework.batch.item.file.mapping.BeanWrapperFieldSetMapper">

<bean:property name="prototypeBeanName" value="DNA\_Sequence">

</bean:property>

</bean:bean>

</bean:property>

</bean:bean>

</bean:property>

</bean:bean>

<!-- lineTokenizer -->

<bean:bean id="lineTokenizer"

class="org.springframework.batch.item.file.transform.DelimitedLineTokenizer">

<bean:property name="delimiter" value=","/>

<bean:property name="names">

<bean:list>

<bean:value>dna</bean:value>

<bean:value>crissprArrays</bean:value>

</bean:list>

</bean:property>

</bean:bean>

<bean:bean id="csvItemWriter"

class="org.springframework.batch.item.file.FlatFileItemWriter"

scope="step">

<bean:property name="resource" value="file:target/ch02/outputFile.csv"/>

<bean:property name="lineAggregator">

<bean:bean

class="org.springframework.batch.item.file.transform.DelimitedLineAggregator">

<bean:property name="delimiter" value="|"></bean:property>

<bean:property name="fieldExtractor">

<bean:bean

class="org.springframework.batch.item.file.transform.BeanWrapperFieldExtractor">

<bean:property name="names"

value="dna,crissprArrays">

</bean:property>

</bean:bean>

</bean:property>

</bean:bean>

</bean:property>

</bean:bean>

I hope you figured out the nuance on how Spring Batch create CRISPR Arrays in outputFile.csv

**The 2nd Phase disabling the viruses another Batch process:**

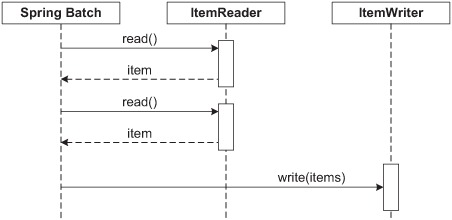
The CRISPR arrays allow the bacteria to "remember" the viruses (or closely related ones). If the viruses attack again,

 the bacteria produce RNA segments from the CRISPR arrays to target the viruses

The bacteria (Spring Batch) then use Cas9 or a similar enzyme to cut the DNA apart, which disables the virus.

 <=> Spring batch write(transformed items) ----> ItemWriter ( cut Virus DNA ).

It is another Batch where the reader reads precedent outputFile.csv (CRISPR arrays) and cut simultaneously the virus.



<step id="step2">

<tasklet transaction-manager="transactionManager">

<chunk reader="itemReader" writer="itemWriter" commit-interval="10"/>

</tasklet>

</step>

**Conclusion**

As You can see in this brief explanation what was done in only one batch step job previously is divided into 2 batch job each job with one step the first batch job step creates the CRISPR Arrays. And the second batch job step is the second phase disabling the viruses simultaneously after reading CRISPR ARRAYS file outputFile.csv and altering the viruses DNA file write(items)

data/DNA.csv

Spring Batch + Karp Rabin = how CRISPR Cas9 works is my IT theoretical model may be it could be interesting and useful for drugs discovery. The model is an idea that had been haunting me since 2012. I share it with you. I can’t go further with it, may be you find it useful interesting and continue developement. The model is under MIT License

<https://github.com/didipostman/CrisprCas9>