replaceAll, replace

- description : 정규표현식을 사용한 replaceall 활용 연구
- author : 주레피
- email : dhan@repia.com
- lastupdate : 2020-04-14

Case study

1. 이미지 소스의 base64문자열 제거 (→)

```
String testStr="<img src=\" ...
5XTk8evT\">";
System.out.println(testStr.replaceAll("src=\"data:image[^\"]*\"",
"src=\"\""));
```

2. http://www.repia.com → www.repia.com(링크추가)

```
String testStr="http://www.repia.com ";
System.out.println(testStr.replaceAll("(http(s)?://([^ ]*))", "<a
href=\"$1\" target=\"blank\">$3</a>"));
```

> 괄호 순서대로
> (http(s)?://([^]*)) ⇒ $1
> (s) ⇒ $2
> ([^]*) ⇒ $3

```
// Source
public static void main(String[] args)
{
    String[] linkHostList = {
"http://www.sit.re.kr/kr/board/result/boardView.do?bbsIdx="
                            ,
"https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn="
                            ,
"https://www.kosen21.org/info/kosenReport/reportView.do?articleSeq=" };

    for(String linkHost: linkHostList) {
        System.out.printf("%s -> %s\n", linkHost,
linkHost.replaceAll("(http(s)?://([^/]*)(/.*))", "$3"));
    }
}
// Output <= $3
http://www.sit.re.kr/kr/board/result/boardView.do?bbsIdx= -> www.sit.re.kr
```

```
https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn= ->
scienceon.kisti.re.kr
https://www.kosen21.org/info/kosenReport/reportView.do?articleSeq= ->
www.kosen21.org

// Output <= $1
http://www.sit.re.kr/kr/board/result/boardView.do?bbsIdx= ->
http://www.sit.re.kr/kr/board/result/boardView.do?bbsIdx=
https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn= ->
https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn=
https://www.kosen21.org/info/kosenReport/reportView.do?articleSeq= ->
https://www.kosen21.org/info/kosenReport/reportView.do?articleSeq=

// Output <= $2
http://www.sit.re.kr/kr/board/result/boardView.do?bbsIdx= ->
https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn= -> s
https://www.kosen21.org/info/kosenReport/reportView.do?articleSeq= -> s

// Output <= $4
http://www.sit.re.kr/kr/board/result/boardView.do?bbsIdx= ->
/kr/board/result/boardView.do?bbsIdx=
https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn= ->
/srch/selectPORSrchTrend.do?cn=
https://www.kosen21.org/info/kosenReport/reportView.do?articleSeq= ->
/info/kosenReport/reportView.do?articleSeq=
```

3. 한줄로 간단하게 ltrim, rtrim

```
String testStr=str.replaceAll("^\\s+",, ""); //ltrim
String testStr=str.replaceAll("\\s+$",, ""); //rtrim
```

> '^' mean 처음부터 매칭
> '$' mean 라인마지막 매칭
> \\s+ mean white space

4. 특수 기호 사용하기
4.1 []로 싸주면 문자자체를 인식하는 것

```
* => [*]
+ => [+]
$ => [$]
| => [|]
? => [?]
```

4.2 \\를 붙여줘야 하는 것

```
( -> \\(
) -> \\)
{ -> \\{
```

```
} -> \\}
^ -> \\^
[ -> \\[
] -> \\]
```

4.3 나머지 기호 !#%&@':;-.< >,~` 는 괜찮은 것 같다.

```
String c = "?$(){}*+^|[]";
c = c.replaceAll("[?][$]\\(\\)\\{\\}[*][+]\\^[|]\\[\\]", "");
```

Ref

- Pattern, Regular Experssion
- Regular 온라인 Test
- replaceAll에서 정규표현식 특수문자 취급 자바

Tip

Troubleshooting

주레피, replaceAll, replace

From:
http://125.132.25.164/dokuwiki/ - **문제를 잘 정의하는 것은 문제를 절반 해결한 것이다. - 2023.12**

Permanent link:
http://125.132.25.164/dokuwiki/doku.php?id=wiki:java:string:replaceall&rev=1664930763

Last update: **2022/10/05 09:46**