

# JSON Input

# Issues with JSON

Input split N

```
{
  "created_at" : "Thu, 29 Dec 2011 21:46:01 +0000",
  "from_user" : "xxx",
  "text" : "Lorem ipsum dolor sit amet",
  "children" : [
    {
      "created_at": "Thu, 29 Dec 2011 21:46:01 +0000",
      "username": "yyy"
    },
    {
      "created_at": "Thu, 29 Dec 2011 21:46:01 +0000",
      "username": "zzz"
    }
  ]
},
{
  "created_at" : "Mon, 26 Dec 2011 21:18:37 +0000",
  "from_user" : "xxx",
  "text" : "consectetur adipisicing elit",
  "children" : [
    {
      "created_at": "Thu, 29 Dec 2011 21:46:01 +0000"
    },
    {
      "created_at": "Thu, 29 Dec 2011 21:46:01 +0000"
    }
  ]
}
```

# Solution: One JSON object per line

```
{"created_at" : "Thu, 29 Dec 2011 21:46:01 +0000", "from_user" : ...  
{"created_at" : "Mon, 26 Dec 2011 21:18:37 +0000", "from_user" : ...
```

Classes to access JSON in Java

<https://github.com/douglascrockford/JSON-java>

# Read JSON inside Mapper

```
public class StubMapper extends Mapper<LongWritable, Text, Text, Text> {
```

```
    @Override
```

```
    public void map(LongWritable key, Text value, Context context)
        throws IOException, InterruptedException {
```

```
        String author;
```

```
        String book;
```

```
        String line = value.toString();
```

```
        String[] tuple = line.split("\\n");
```

```
        try{
```

```
            for(int i=0;i<tuple.length; i++){
```

```
                JSONObject obj = new JSONObject(tuple[i]);
```

```
                author = obj.getString("author");
```

```
                book = obj.getString("book");
```

```
                context.write(new Text(author), new Text(book));
```

```
            }
```

```
        }catch(JSONException e){
```

```
            e.printStackTrace();
```

```
        }
```

```
    }
```

```
}
```