

Hide Responsibility in Archive

```
class S3Archive(object):
    def __init__(self, bucket_name, prefix, client):
        """Ctor"""
        self.client = boto3.client('s3', region_name=region_name)
        self.bucket_name = bucket_name
        self.prefix = prefix
```

```
    def items(self):
        for key in self.get_keys():
            lines = self.get_file(key)
            yield key, lines
```

```
    def get_keys(self):
        """Get an iterable of keys"""
        for page in self.get_pages():
            for key in self.keys_from(page):
                yield key
```

```
+---- 7 lines: def get_file(self, key): -----
+---- 4 lines: def get_pages(self): -----
+---- 6 lines: def keys_from(self, page): -----
```

Archive as a Tuple Stream

```
def find_bad_articles(archive, predicate):  
    for key, lines in archive.items():  
        if predicate(lines):  
            yield key
```