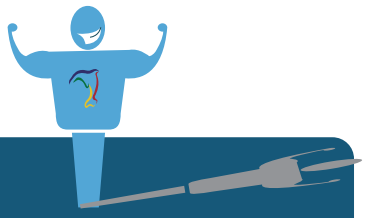
autentia

Avenida de Castilla,1 - Edificio Best Point - Oficina 21B
28830 San Fernando de Henares (Madrid)
tel./fax: +34 91 675 33 06
info@autentia.com - **www.autentia.com**

¿**Qué ofrece** Autentia Real Business Solutions S.L?

Somos su empresa de **Soporte a Desarrollo Informático**.
Ese apoyo que siempre quiso tener...

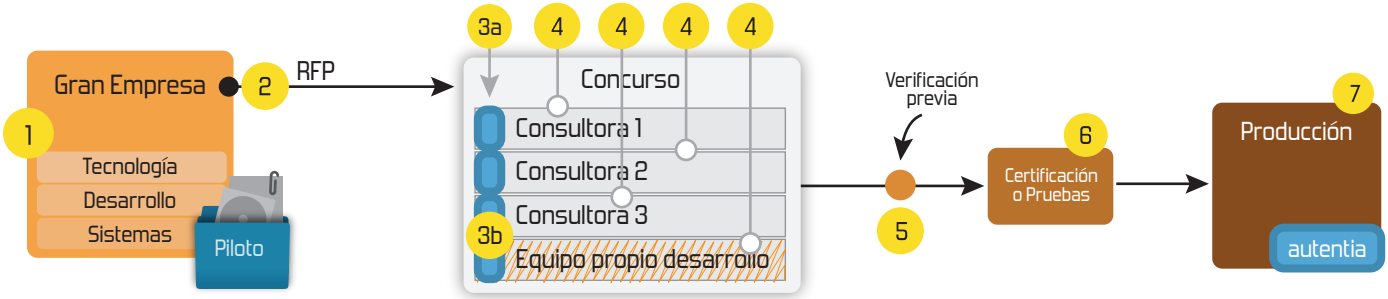
1. Desarrollo de componentes y proyectos a medida



2. Auditoría de código y recomendaciones de mejora

3. Arranque de proyectos basados en nuevas tecnologías

1. Definición de frameworks corporativos.
2. Transferencia de conocimiento de nuevas arquitecturas.
3. Soporte al arranque de proyectos.
4. Auditoría preventiva periódica de calidad.
5. Revisión previa a la certificación de proyectos.
6. Extensión de capacidad de equipos de calidad.
7. Identificación de problemas en producción.

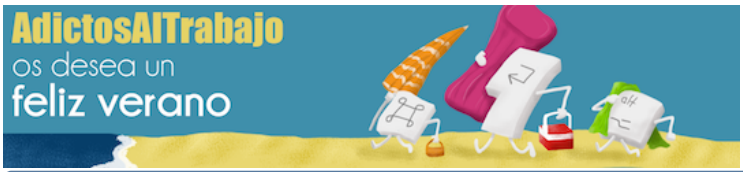


4. Cursos de formación (impartidos por desarrolladores en activo)

- Spring MVC, JSF-PrimeFaces /RichFaces, HTML5, CSS3, JavaScript-jQuery
- Gestor portales (Liferay)
- Gestor de contenidos (Alfresco)
- Aplicaciones híbridas
- Control de autenticación y acceso (Spring Security)
- UDDI
- Web Services
- Rest Services
- Social SSO
- SSO (Cas)
- JPA-Hibernate, MyBatis
- Motor de búsqueda empresarial (Solr)
- ETL (Talend)
- Dirección de Proyectos Informáticos.
- Metodologías ágiles
- Patrones de diseño
- TDD
- Tareas programadas (Quartz)
- Gestor documental (Alfresco)
- Inversión de control (Spring)
- BPM (jBPM o Bonita)
- Generación de informes (JasperReport)
- ESB (Open ESB)

Compartimos nuestro conocimiento en:
www.adictosaltrabajo.com

Para más información visítenos en:
www.autentia.com



Entra en Adictos a través de
E-mail
Contraseña
Entrar
Registrarme
Olvidé mi contraseña

Inicio | Quiénes somos | Formación | Comparador de salarios | Nuestros libros | Más

» Estás en: Inicio » Tutoriales » Merge de ficheros con Talend

Alfonso Blanco Criado

Consultor tecnológico de desarrollo de proyectos informáticos.

Ingeniero Informático *

Creador del blog Rueda Lenticular

Charla sobre Pluto

Puedes encontrarme en Autentia: Ofrecemos servicios de soporte a desarrollo, factoría y formación

Somos expertos en Java/J2EE

Ver todos los tutoriales del autor

Fecha de publicación del tutorial: 2014-10-16

Tutorial visitado 1 veces Descargar en PDF

Merge de ficheros con Talend

0. Indice de contenidos.

- 1. Introducción
- 2. Entorno
- 3. Creación y configuración
- 4. Ejecución
- 5. Conclusiones

1. Introducción

Talend es una herramienta ETL muy extendida dentro de un entorno de trabajo. Cada vez que se necesita hacer ciertos procesos de carga de datos con una cierta periodicidad o bien sean puntuales pero con una complejidad se piensa en este tipo de herramientas. Talend es una herramienta muy potente y bastante *"fácil"* de usar por lo que en Autentia forma parte de nuestro ecosistema de desarrollo desde hace tiempo y se nota en la cantidad de tutoriales que tenemos ya publicados.

En el tutorial de hoy nos vamos a centrar en el componente **tUnite** que nos permitirá unir/mergear varios ficheros/fuentes de datos en uno sólo.

2. Entorno.

El tutorial está escrito usando el siguiente entorno:

- Hardware: Portátil MacBook Pro 15' (2.5 GHz Intel Core i7, 16GB DDR3).
- Sistema Operativo: Mac OS Mavericks 10.9.5
- Talend Open Studio 5.5.1

3. Creación y configuración

Nos creamos un nuevo job dentro de nuestro proyecto de Talend e incorporamos a él dos componentes para la lectura de ficheros del tipo **tFileInputFullRow**. Los configuramos para que lean dos ficheros de texto diferentes, uno con la cabecera y otro con los datos. A continuación incorporamos un componente de tipo **tUnite** que se encargará de hacer el merge o fusión de las dos fuentes de datos. Hay que tener en cuenta el orden para que el primero sea el componente que lee el fichero de cabecera y el segundo el componente que lee el fichero de datos. Este orden se establece vinculando primero el componente tFileInputFullRow de la cabecera al componente tUnite y posteriormente el componente tFileInputFullRow de los datos.

Una de las particularidades del componente tUnite es que todas las fuentes de datos entrantes a él tienen que tener un esquema común, en nuestro caso será una cadena de texto que representa una línea leída del fichero. Además el componente tUnite tiene una variable global llamada *"Number of lines"* que nos indicará el número de líneas procesadas por él tras su ejecución.

Para finalizar hemos incorporado otro componente de tipo tFileOutputDelimited para poder comprobar el resultado de nuestro job, aunque este paso es opcional. Es decir, se podría haber puesto un componente de log para poder verlo u otro componente de *"salida de datos"*.

Catálogo de servicios Autentia

Síguenos a través de:

Últimas Noticias

» Curso JBoss de Red Hat

» Si eres el responsable o líder técnico, considérate desafortunado. No puedes culpar a nadie por ser gris

» Portales, gestores de contenidos documentales y desarrollos a medida

» Comentando el libro Start-up Nation, La historia del milagro económico de Israel, de Dan Senor & Salu Singer

» Screencasts de programación narrados en Español

Histórico de noticias

Últimos Tutoriales

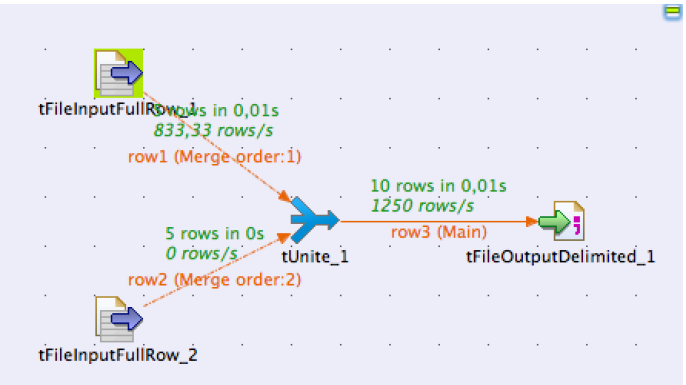
» Jugando con AngularJS y Google Maps

» Extendiendo las reglas de SonarQube con Xpath

» Primeros pasos con Apache Kafka

» Solución de problemas comunes con la integración de maven en Eclipse Luna.

» Integración de SonarQube



4. Ejecución

Una vez con todos los componentes incorporados a nuestro job, configurados y enlazados entre sí, solo nos queda ejecutar el job y comprobar su funcionamiento. Vemos tras su ejecución que el fichero de salida generado es la unión, en el orden deseado, de los dos ficheros de entrada.

```
Cabecera del fichero
--------------------
Tutorial de AdictosAlTrabajo.com
Merge de ficheros con Talend

```

Fichero de cabecera

```
Primera línea
Segunda línea
Tercera línea
Cuarta línea
Quinta línea
```

Fichero de datos

```
Cabecera del fichero
--------------------
Tutorial de AdictosAlTrabajo.com
Merge de ficheros con Talend

Primera línea
Segunda línea
Tercera línea
Cuarta línea
Quinta línea
```

Fichero final con la unión de ambos ficheros

5. Conclusiones

Como se puede ver el uso de componentes como el que hemos visto hoy, tUnite, es muy sencillo y nos puede ayudar mucho a la hora de procesar ficheros, sobre todo como parte de un proceso más complejo.

Espero que este tutorial os haya sido de ayuda. Un saludo.

Alfonso Blanco Criado ablanco@autentia.com

en Eclipse.

Últimos Tutoriales del Autor

» Código de barras con iReport

» Eventos en MySQL

» Estilos externos en iReport

» Informes Crosstab con iReport

» JasperReports Maven Plugin

A continuación puedes evaluarlo:

Regístrate para evaluarlo

Por favor, vota +1 o compártelo si te pareció interesante

| Share

Anímate y coméntanos lo que pienses sobre este **TUTORIAL:**

» Registrate y accede a esta y otras ventajas **«**

SOME RIGHTS RESERVED Esta obra está licenciada bajo licencia Creative Commons de Reconocimiento-No comercial-Sin obras derivadas 2.5

Copyright 2003-2014 © All Rights Reserved | Texto legal y condiciones de uso | Banners | Powered by Autentia | Contacto

XHTML 1.0 CSS XML RSS XML ATOM