



Herramientas de Análisis de la Información de la AEAT: ZUJAR



PROBLEMÁTICA DEL ANÁLISIS DE INFORMACIÓN EN LA AEAT

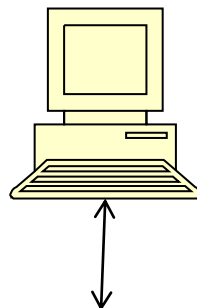
Tres factores que hacen difícil encontrar herramientas comerciales que den un soporte integrado a toda la organización:

- La abundancia de datos, existe un **gran número de entidades** a analizar (Contribuyentes, Expedientes, Vehículos, Cuentas Bancarias, cientos de Modelos Tributarios, ...)
- Un **gran número de registros** por entidad (en algunos casos puede llegar a miles de millones de registros).
- **Gran número de variables** que se pueden analizar en el caso concreto de la información tributaria asociada al Contribuyente (miles de variables), con origen en declaraciones y autoliquidaciones que de forma periódica deben presentar los obligados tributarios, así como de otra que es requerida por la Inspección durante la instrucción de los expedientes, bien a los interesados o a terceros.

DESARROLLO PROPIO

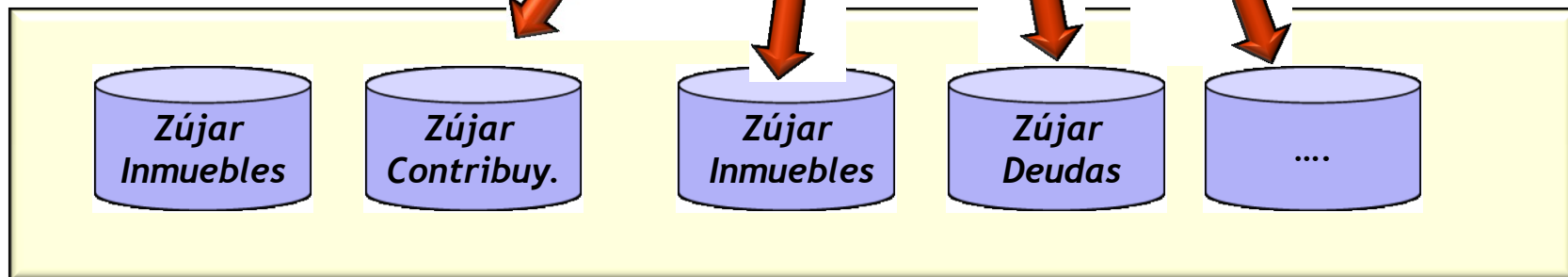
Arquitectura de la herramienta Zújar

APLICACIÓN CLIENTE ZUJAR



MOTOR ZUJAR (SERVIDOR)

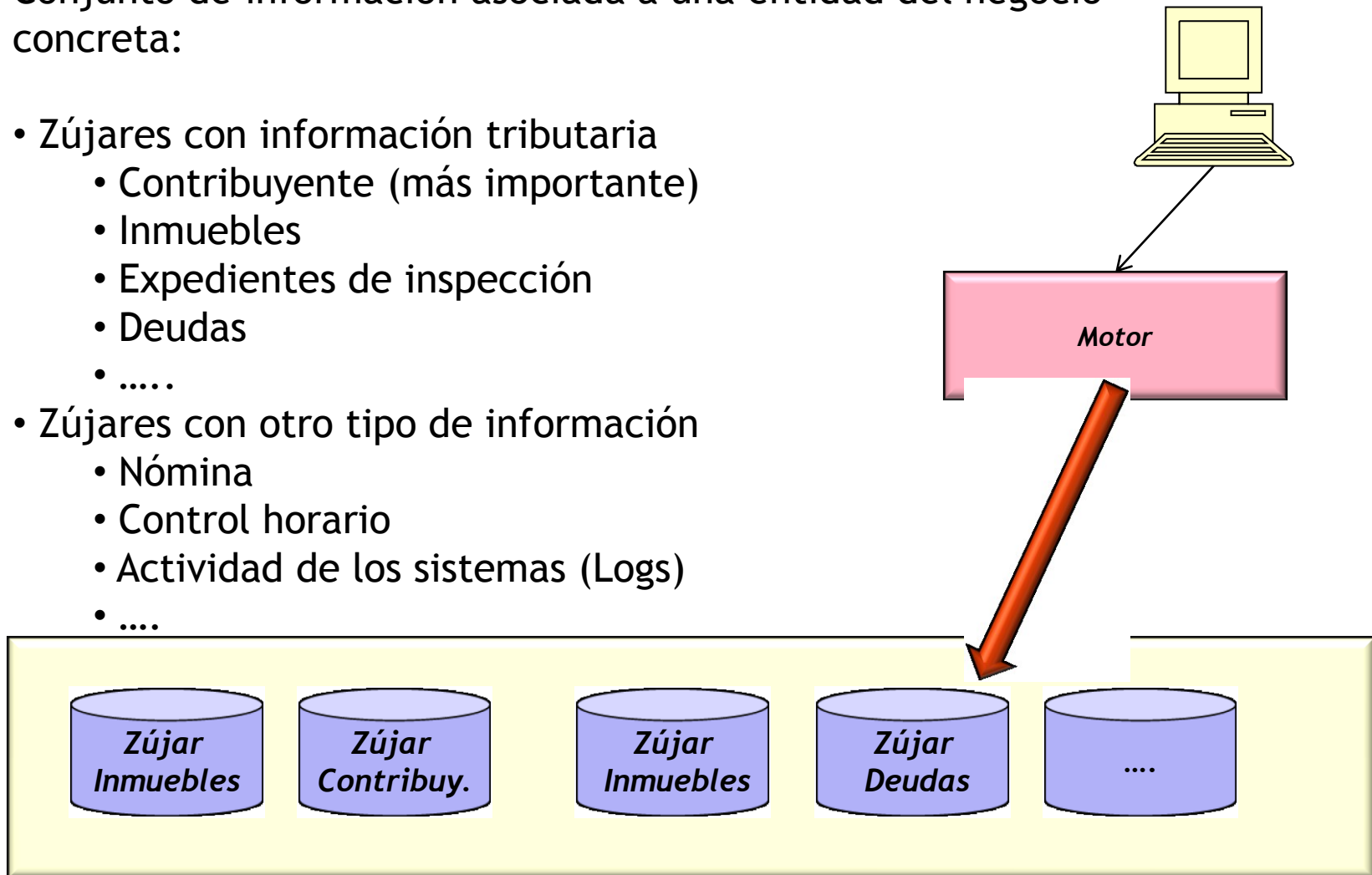
BASE DE DATOS



¿Qué es una entidad Zújar?

Conjunto de información asociada a una entidad del negocio concreta:

- Zújares con información tributaria
 - Contribuyente (más importante)
 - Inmuebles
 - Expedientes de inspección
 - Deudas
 -
- Zújares con otro tipo de información
 - Nómina
 - Control horario
 - Actividad de los sistemas (Logs)
 -



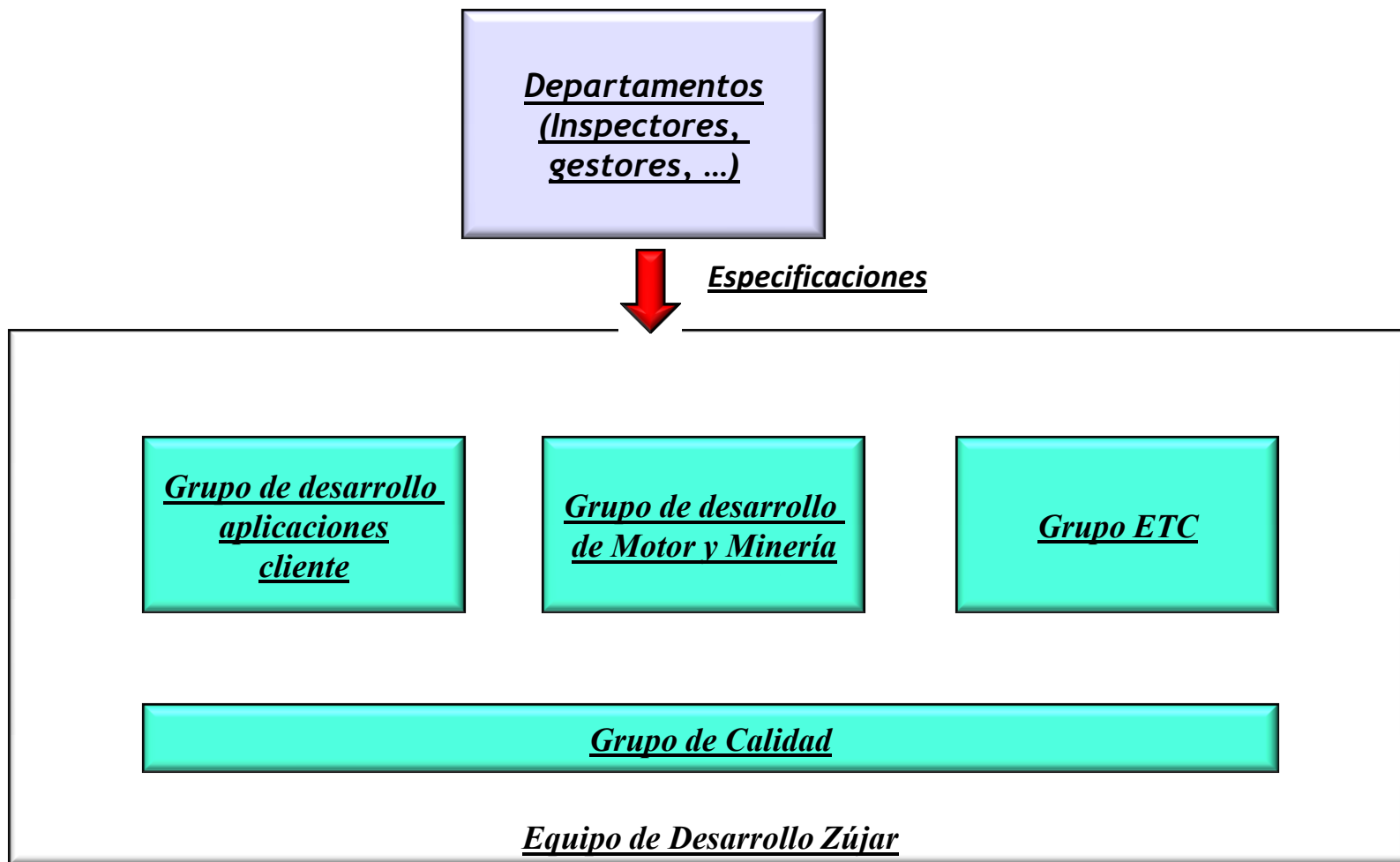
Propósito de la herramienta Zújar

- **Herramientas de Business Intelligence (BI) o Inteligencia de Negocio** para análisis multidimensional de la información ajustada a las necesidades de la AEAT y orientada a la toma de decisión y mejora de los objetivos de la organización.
- **Infraestructura de acceso potente a datos** para otras aplicaciones.
- **Herramienta de selección (contribuyentes):**
 - filtrado sobre miles de variables y sobre millones de registros.
 - proceso **interactivo** (tiempos de respuesta de pocos segundos)

Para generar el colectivo de contribuyentes sobre los que se van a tomar determinadas acciones tanto de gestión como de inspección, ahorrando tiempo al usuario analizando contribuyentes que no corresponden al colectivo a tratar.

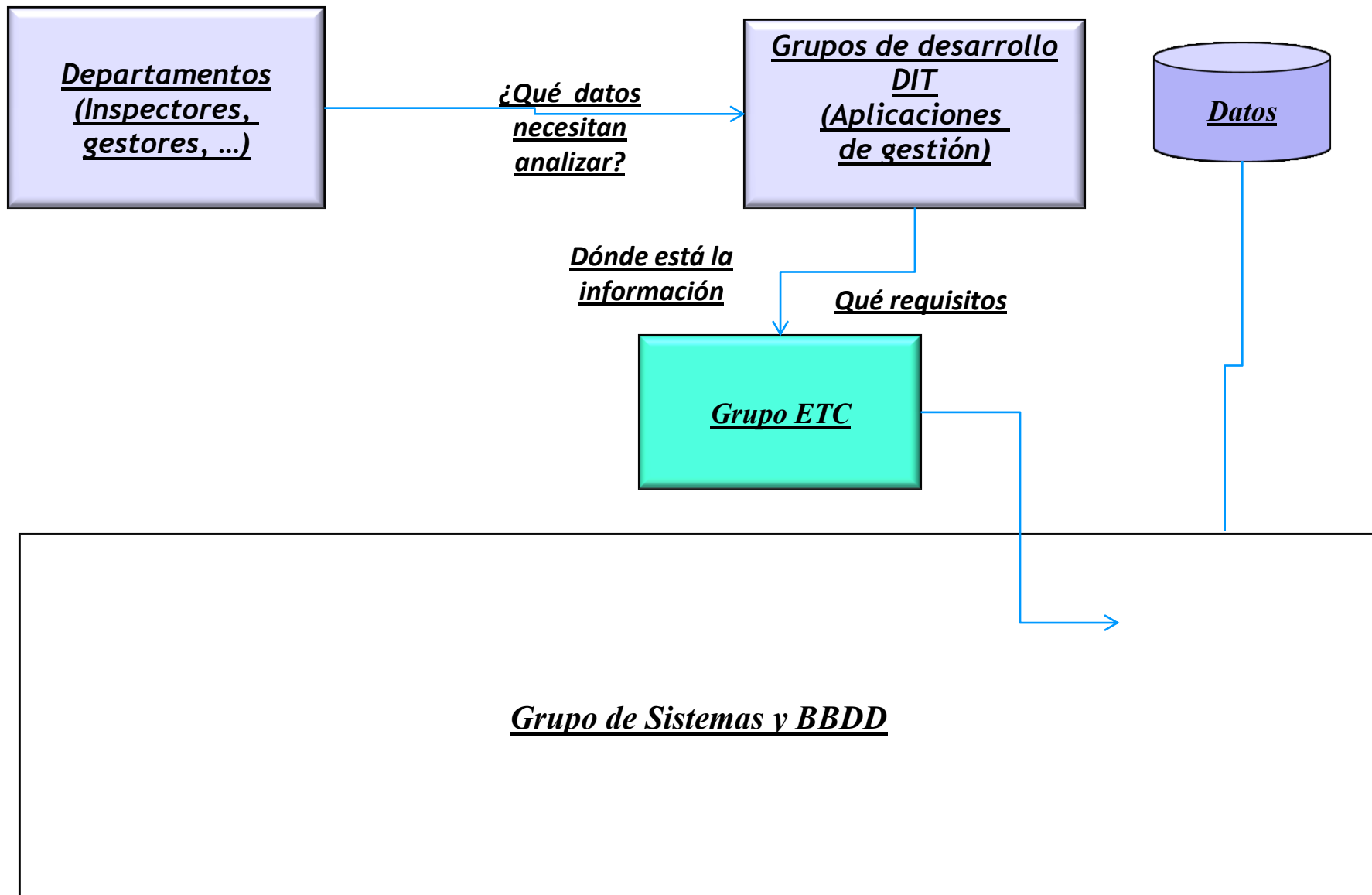


EQUIPOS PARTICIPANTES EN EL DESARROLLO



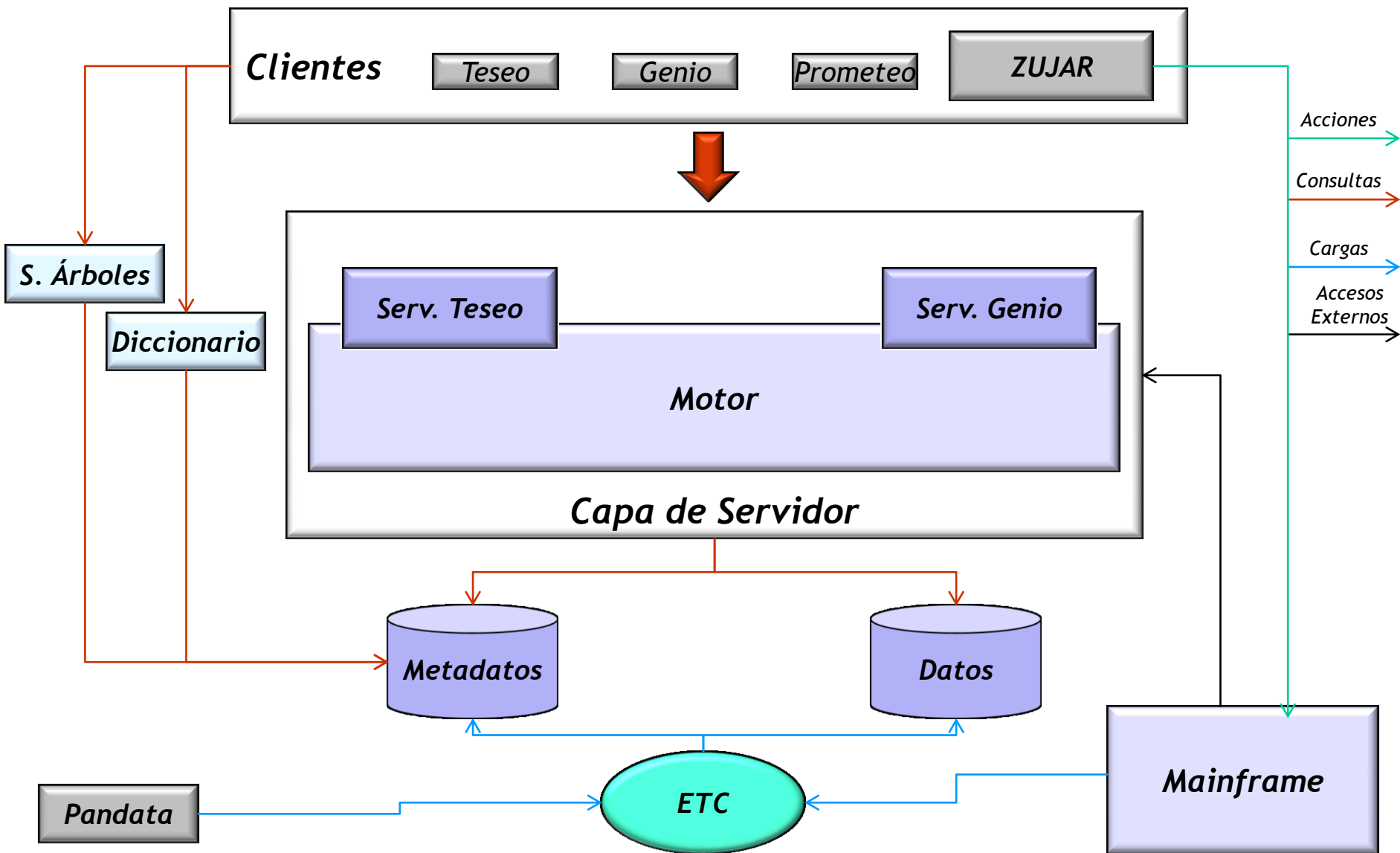


Participantes en la elaboración de una entidad Zújar





Arquitectura Lógica





Interfaz de usuario (Zújar)

Contribuyentes - MultiZujar 9.0.0.12

Inicio Menú Configuración / Ayuda

Página principal PETICIONES DE USUARIO Información Externa LOG HTTPSERVERS INTERNET SEPRI LOG Httpservers Internet Histórico Septiembre 2014 Contribuyentes [2014] X

Multiregistro* X

Autofiltro Condiciones Aplicar Ejercicios QD Columnas Cancelar

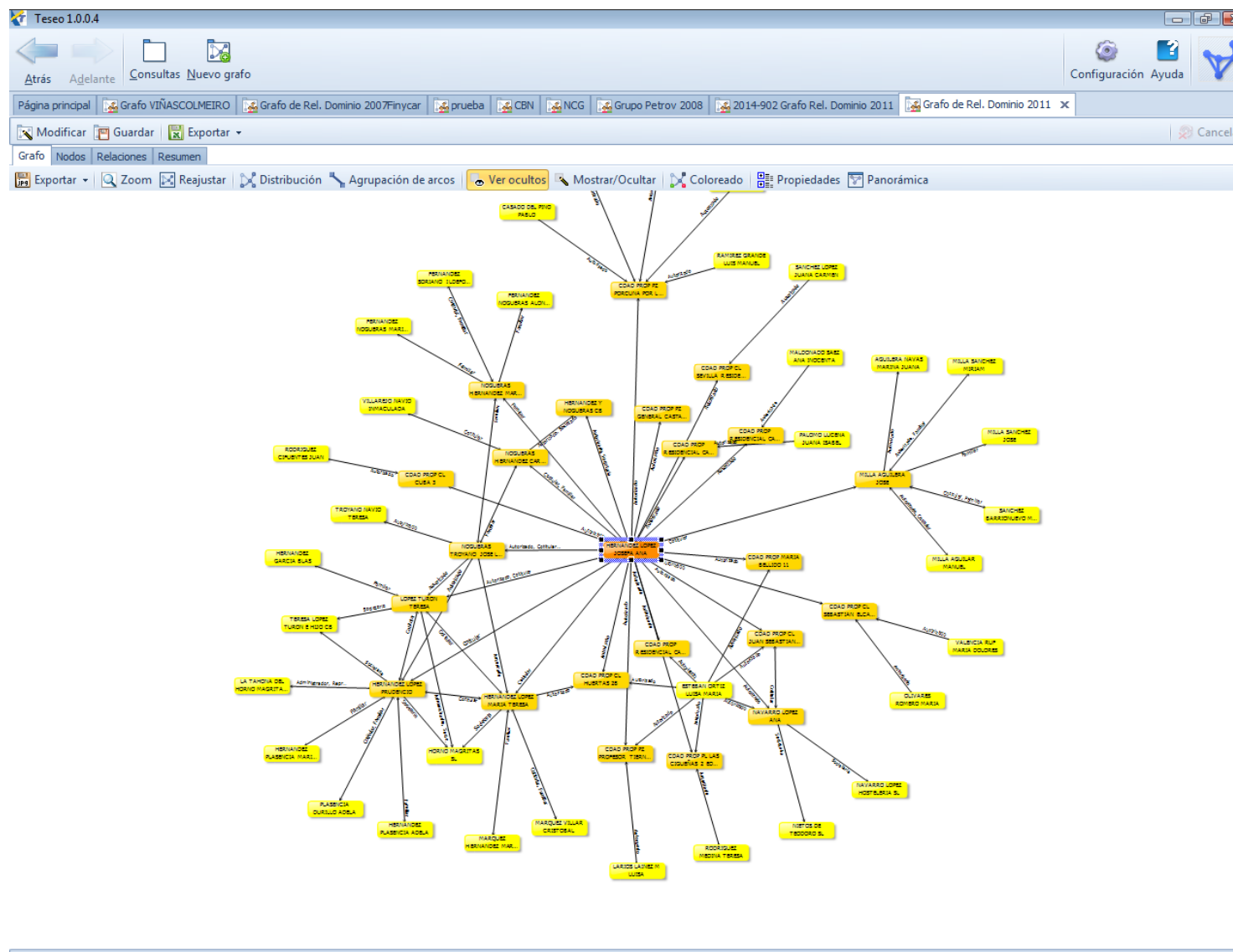
[V.Total] (2014)Total adquisición vehículos mayor o igual que introduzca un valor

[Domicilio Fiscal] Literal de (2014)Delegación Especial	(2014)Fecha nacimiento	[Domicilio Fiscal] Literal de (2014)Delegación	[Domicilio Fiscal] Literal de (2014)Administración	[Domicilio Fiscal] Literal de (2014)Municipio	[Domicilio Fiscal] (2014)Nombre de vía	Literal de (2014)Tipo de NIF	[Estado del Cte.] Literal de (2014)Cte. operativo
País Vasco	10-12-1952	Álava	Álava	Llodio	LARRA O	Persona Física	Alta
País Vasco	30-12-1954	Álava	Álava	Amurrio	LANDAKO	Persona Física	Alta
País Vasco	19-06-1955	Álava	Álava	Llodio	GOIKOGANE	Persona Física	Alta
País Vasco	30-04-1953	Álava	Álava	Llodio	ZUMALACARREGUI	Persona Física	Alta
País Vasco	30-04-1957	Álava	Álava	Artziniega	EGUZQUIMENDI BQ E	Persona Física	Alta
País Vasco	05-10-1957	Álava	Álava	Vitoria-Gasteiz	ESPAÑA PLAZA DE	Persona Física	Alta
País Vasco	12-12-1955	Álava	Álava	Llodio	ALBERTO ACERO	Persona Física	Alta
Castilla y León	26-10-1955	Burgos	Burgos	Merindad de Cuesta-Urria	LA FUENTE	Persona Física	Alta
País Vasco	08-09-1956	Álava	Álava	Amurrio	IRITXUSI	Persona Física	Baja
País Vasco	10-05-1956	Álava	Álava	Llodio	SANTA MARIA DEL YERMO 4	Persona Física	Alta
País Vasco	01-06-1957	Álava	Álava	Vitoria-Gasteiz	ERREKATXIKI	Persona Física	Alta
País Vasco	08-06-1956	Álava	Álava	Vitoria-Gasteiz	GOMECHA	Persona Física	Alta
País Vasco	05-06-1956	Álava	Álava	Amurrio	LANDAKO	Persona Física	Alta
País Vasco	17-06-1956	Álava	Álava	Llodio	MAESTRO ELORZA	Persona Física	Alta
País Vasco	07-07-1956	Álava	Álava	Amurrio	MENDIKO	Persona Física	Alta
País Vasco	16-01-1956	Álava	Álava	Llodio	ESKUTXI	Persona Física	Alta
País Vasco	18-10-1957	Álava	Álava	Llodio	LANDALUZE	Persona Física	Baja
País Vasco	13-11-1956	Álava	Álava	Amurrio	ETXEGOIEN	Persona Física	Alta
País Vasco	27-06-1954	Álava	Álava	Llodio	VIRGEN DEL CARMEN	Persona Física	Alta
País Vasco	29-10-1956	Álava	Álava	Llodio	ZUMALACARREGUI	Persona Física	Alta
País Vasco	29-08-1957	Vizcaya	Vizcaya	Arrigorriaga	URGOITI	Persona Física	Alta
País Vasco	07-04-1957	Álava	Álava	Llodio	UGARTE	Persona Física	Alta
País Vasco	23-07-1957	Álava	Álava	Llodio	OKETA	Persona Física	Alta
Castilla y León	12-02-1963	Burgos	Burgos	Miranda de Ebro	CARLOS III	Persona Física	Alta
País Vasco		Álava	Álava	Llodio	PAGOLAR	Persona Física	Baja
País Vasco	15-10-1959	Vizcaya	Vizcaya	Bilbao	COSME ECHEVARRIETA	Persona Física	Alta
País Vasco	20-04-1959	Álava	Álava	Amurrio	DIONISIO ALDAMA	Persona Física	Alta
País Vasco	05-01-1959	Álava	Álava	Llodio	ZUMALACARREGUI	Persona Física	Alta
País Vasco	05-07-1959	Álava	Álava	Llodio	LANDALUZE	Persona Física	Alta
País Vasco	07-01-1960	Álava	Álava	Llodio	LARRA O	Persona Física	Alta
País Vasco	05-05-1960	Álava	Álava	Amurrio	GOIKOLARRA	Persona Física	Alta
País Vasco	15-11-1958	Álava	Álava	Vitoria-Gasteiz	LANDAZURI JOSE J	Persona Física	Alta
País Vasco		Álava	Álava	Llodio	NERVION 3 S D	Persona Física	Alta
País Vasco	11-05-1961	Álava	Álava	Amurrio	JOSE PICAZA	Persona Física	Alta
País Vasco	30-05-1961	Álava	Álava	Amurrio	LARRIMBE S N	Persona Física	Alta

Se muestran los 1.000 primeros contribuyentes de un total de 65.424.393.

8

Interfaz de usuario (TESEO)



Genio: generador de informes

4.- Información relacionada con ingresos y pagos

4.1.- Ingresos y pagos

Origen de la información: Modelo 347.

A partir del modelo de declaración informativa número 347 presentado por el propio contribuyente (declarado) o por terceros que operaron con él (imputados), se obtienen las siguientes cifras de ingresos y pagos:

Año	INGRESOS				PAGOS			
	Declarados		Imputados		Declarados		Imputados	
	Importe	Núm.	Importe	Núm.	Importe	Núm.	Importe	Núm.
2014	0,00	0	10.520.119.290,67	144	0,00	0	188.181.315,56	1.065
2013	0,00	0	7.539.629.455,07	168	0,00	0	178.393.911,82	1.075
2012	0,00	0	6.563.755.747,30	217	0,00	0	200.244.079,92	1.120
2011	20.411,76	1	6.713.127.718,56	222	3.140.478,67	1	235.939.681,11	1.168
2010	139.617,12	2	7.227.144.990,38	216	3.146.678,44	1	262.382.132,68	1.167
2009	26.382,24	2	7.454.954.212,73	227	4.388.837,32	2	316.007.170,87	1.176

4.2.- IVA: Base imponible y volumen de operaciones

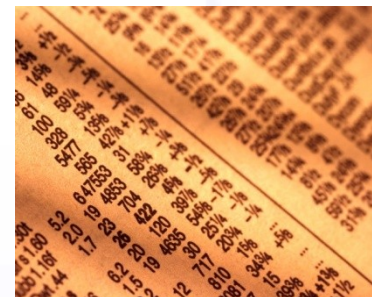
Origen de la información: Modelo 390.

A partir del modelo 390, Resumen Anual de IVA, el contribuyente declara las bases imponibles y volúmenes de operaciones siguientes:

IVA		
Ejercicio	Base imponible	Volumen de operaciones
2014	42.227.909,58	554.187.102,89
2013	46.340.577,04	519.157.929,22
2012	19.819.368,09	543.136.117,33
2011	16.608.434,45	559.606.537,04

Modelo de datos

- Tablón (un Zújar una tabla):
 - Desnormalización máxima (< 3.000 columnas)
 - Minimizar número de cruces
 - Columnas precalculadas para optimización de consultas
- Estrella (un Zújar múltiples tablas diferentes estructuras):
 - En realidad es una particionado vertical (> 3.000 columnas)
 - Relaciones 1:1
 - Joins por clave única: Contribuyente (NIU)
- Particionado horizontal (un Zújar múltiples tablas misma estructura):
 - Tablas con acceso acotado por una variable (LOG - Año)





Modelo de datos: particionado horizontal

Tablas con miles de millones de registros

Variable División	Tabla Física
V1	tabla1
V2	tabla1
V3	tabla2
V4	tabla3
V5	tabla3

Tabla Índice
(Año, mes,
NIU,)

Zújar total



Variable División	Atrib1	Atrib2
V1		
V1		
...		
V1		
V2		
...		
V2		

tabla1

Variable División	Atrib1	Atrib2
V3		
V3		
V3		
...		
V3		

Zújares parciales

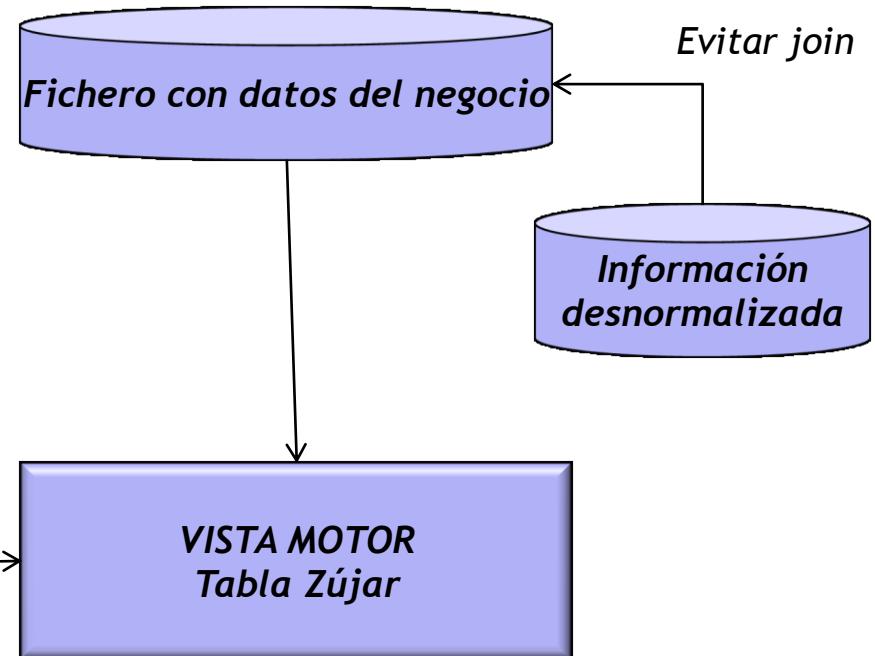
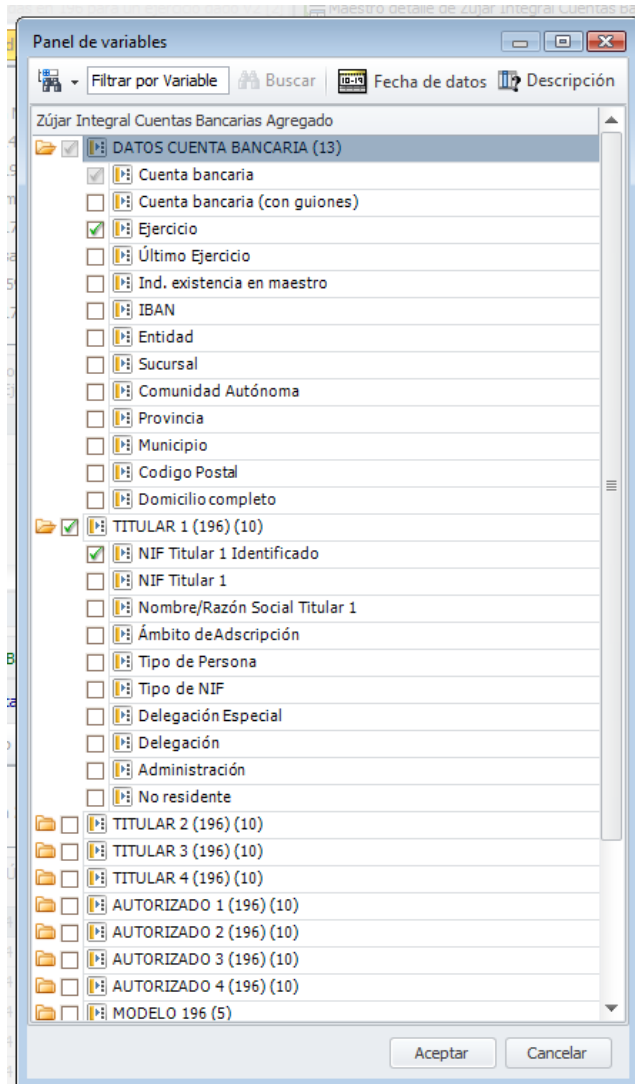
tabla2

Variable División	Atrib1	Atrib2
V4		
V4		
....		
V4		
V5		
...		
V5		

tabla3

Modelo de datos: diccionario Zújar

VISTA CLIENTE: árbol de variables



DICCIONARIO DE METADATOS: recoge información de ficheros, tablas, árboles de variables (vista de usuario), seguridad, relaciones entre Zújares, ...

Funcionalidades Zújar

FILTRADO DE REGISTROS

- Operadores lógicos: Y, O, NOT, ...
- Operadores de relación (=, >, <,)
- Operadores especiales (colectivos)

VISTA DE RESULTADOS

- Tabla
- Gráficos
- Vistas enlazadas
- Exportaciones (Excel, Access, ...)

OPERACIONES SOBRE LOS RESULTADOS

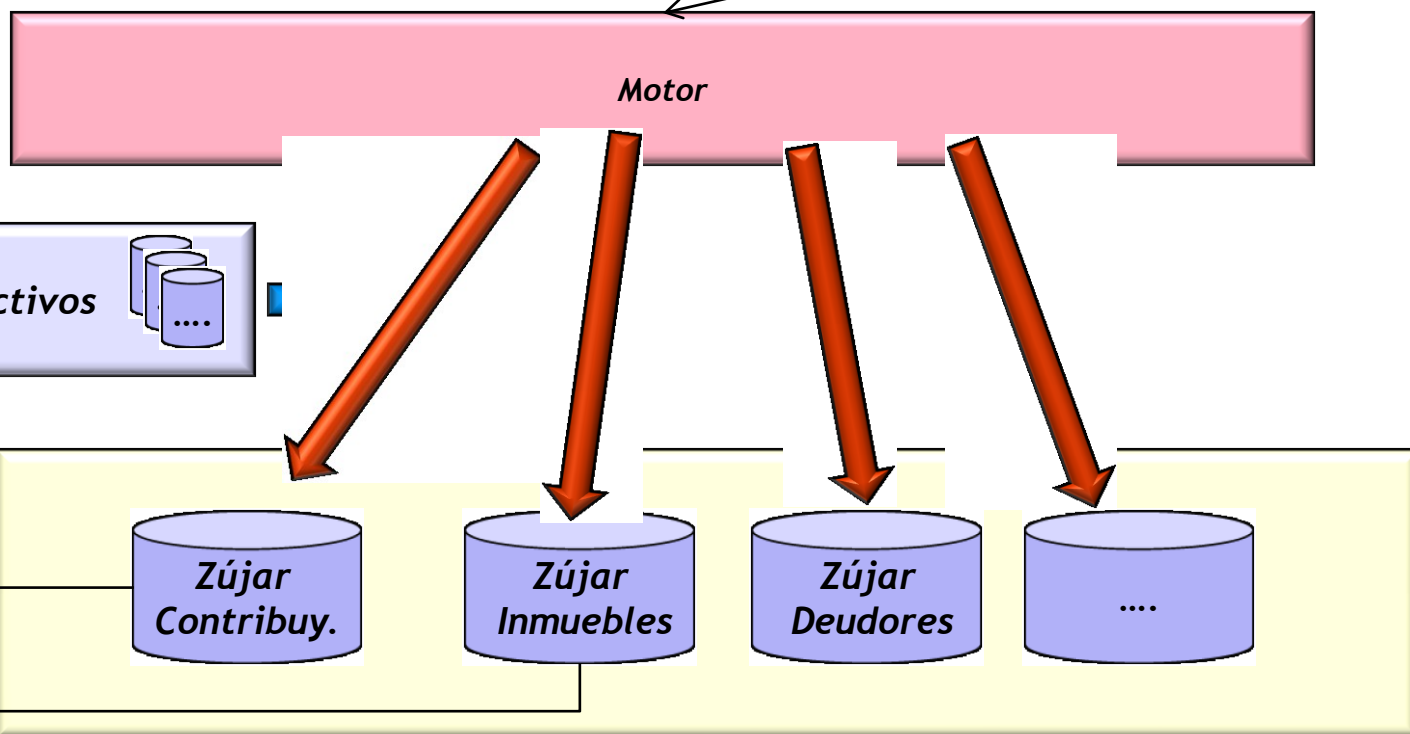
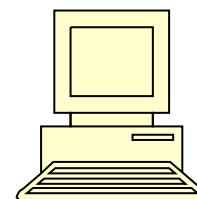
- Ordenación
- Estadísticos (Suma, Media, ...)
- Agrupaciones
- Expresiones calculadas ($v1+v2/v3...$)
- Consultas (publicas y privadas)
- Reglas de formato (resaltar los registros que ...)

Cruce de peticiones entre Zújares

Colectivos creados en otros Zújares se pueden aplicar como filtro.

Enlaces en cliente (local) entre Zújares.

Combinación de Zújares en próxima versión.



Cruce de peticiones entre Zújares: colectivos

NIF Contribuyente Identificado pertenece a colectivo **“mi colectivo”**

(2010) Tipo de NIF igual a introduzca un valor

Total Declarado 2010 mayor o igual que 700.000,00

Total imputado 2010 mayor o igual que 700.000,00

[Cuentas Bancarias] (2010) Saldo último trimestre menor o igual que 3.000,00

[Cuentas Bancarias] (2010) Saldo a 31 de diciembre menor o igual que 3.000,00

[Subconcepto] (2010) Relación Partícipes M.036 igual a Sí

[Subconcepto] (2010) Representados M.036 igual a Sí

[P.Hip.] (2010) Importe préstamo (imputado) menor o igual que 0,00

[P.Hip.] (2010) Importe préstamo (imputado) es nulo

FILTROS

Contribuyente
PERTENECE A COLECTIVO
“mi colectivo” (de Zújar 1)

(2010) Tipo de NIF	[Estado del Cte.] (2014) Fecha alta BDC	[Estado del Cte.] (2010) Fecha alta BDC	[Total x Conceptos] (2010) Saldo total autoliquidaciones
► Sdad. Resp. Limitada	13-07-2009	13-07-2009	1.037,35
Sdad. Resp. Limitada	17-03-2004	17-03-2004	
Sdad. Resp. Limitada	23-03-2007	23-03-2007	
Sdad. Resp. Limitada	16-02-2010	16-02-2010	3.032,86
Sdad. Resp. Limitada	14-11-2006	14-11-2006	
Sdad. Resp. Limitada	18-04-2005	18-04-2005	
Sdad. Resp. Limitada	18-04-2006	18-04-2006	4.383,47
Sdad. Resp. Limitada	12-06-2008	12-06-2008	
Sdad. Resp. Limitada	11-11-2008	11-11-2008	5.292,58

RESULTADOS (Zújar 2)

Cruce de peticiones entre Zújares: enlace entre Zújares

	(2010) Tipo de NIF	[Estado del Cte.] (2014) Fecha alta BDC	[Estado del Cte.] (2010) Fecha alta BDC	[Total x Conceptos] (2010) Saldo total autoliquidaciones
►	Sdad. Resp. Limitada	13-07-2009	13-07-2009	1.037,35
	Sdad. Resp. Limitada	17-03-2004	17-03-2004	
	Sdad. Resp. Limitada	23-03-2007	23-03-2007	
	Sdad. Resp. Limitada	16-02-2010	16-02-2010	3.
	Sdad. Resp. Limitada	14-11-2006	14-11-2006	
	Sdad. Resp. Limitada	18-04-2005	18-04-2005	
	Sdad. Resp. Limitada	18-04-2006	18-04-2006	4.383,47
	Sdad. Resp. Limitada	12-06-2008	12-06-2008	
	Sdad. Resp. Limitada	11-11-2008	11-11-2008	5.292,58

RESULTADOS Petición 1 (Zújar 1)

	[Domicilio Fiscal] (2011) Delegación Especial	[Domicilio Fiscal] (2011) Delegación	[Domicilio Fiscal] (2011) Administración
►	Andalucía	Córdoba	Córdoba
	Valencia	Alicante	Elda
	Andalucía	Almería	Almería
	Castilla y León	Ávila	Ávila
	Extremadura	Badajoz	Badajoz
	Baleares	Illes Balears	Inca
	Castilla La Mancha	Ciudad Real	Ciudad Real
	Andalucía	Córdoba	Lucena
	Andalucía	Granada	Granada Almanjajar
	Andalucía	Granada	Granada
	Andalucía	Granada	Granada
	Castilla La Mancha	Toledo	Toledo
	Castilla La Mancha	Toledo	Toledo

RESULTADOS Petición 2 (Zújar 2)

Contribuyente del Zújar 2 que están a la consulta del Zújar 1

Cruce de peticiones entre Zújares: combinación

Peticiones (filtros, ordenaciones, ...) sobre campos de distintos Zújares (siguiente versión)

(2010) Tipo de NIF	[Estado del Cte.] (2014) Fecha alta BDC	[Estado del Cte.] (2010) Fecha alta BDC	[Total x Conceptos] (2010) Saldo total autoliquidaciones
► Sdad. Resp. Limitada	13-07-2009	13-07-2009	1.037,35
Sdad. Resp. Limitada	17-03-2004	17-03-2004	
Sdad. Resp. Limitada	23-03-2007	23-03-2007	
Sdad. Resp. Limitada	16-02-2010	16-02-2010	3.032,86
Sdad. Resp. Limitada	14-11-2006	14-11-2006	
Sdad. Resp. Limitada	18-04-2005	18-04-2005	
Sdad. Resp. Limitada	18-04-2006	18-04-2006	4.383,47
Sdad. Resp. Limitada	12-06-2008	12-06-2008	
Sdad. Resp. Limitada	11-11-2008	11-11-2008	5.292,58

ZUJAR 1

[Domicilio Fiscal] (2011) Delegación Especial	[Domicilio Fiscal] (2011) Delegación	[Domicilio Fiscal] (2011) Administración
► Andalucía	Córdoba	Córdoba
Valencia	Alicante	Elda
Andalucía	Almería	Almería
Castilla y León	Ávila	Ávila
Extremadura	Badajoz	Badajoz
Baleares	Illes Balears	Inca
Castilla La Mancha	Ciudad Real	Ciudad Real
Andalucía	Córdoba	Lucena
Andalucía	Granada	Granada Almanjajar
Andalucía	Granada	Granada
Andalucía	Granada	Granada
Castilla La Mancha	Toledo	Toledo
Castilla La Mancha	Toledo	Toledo

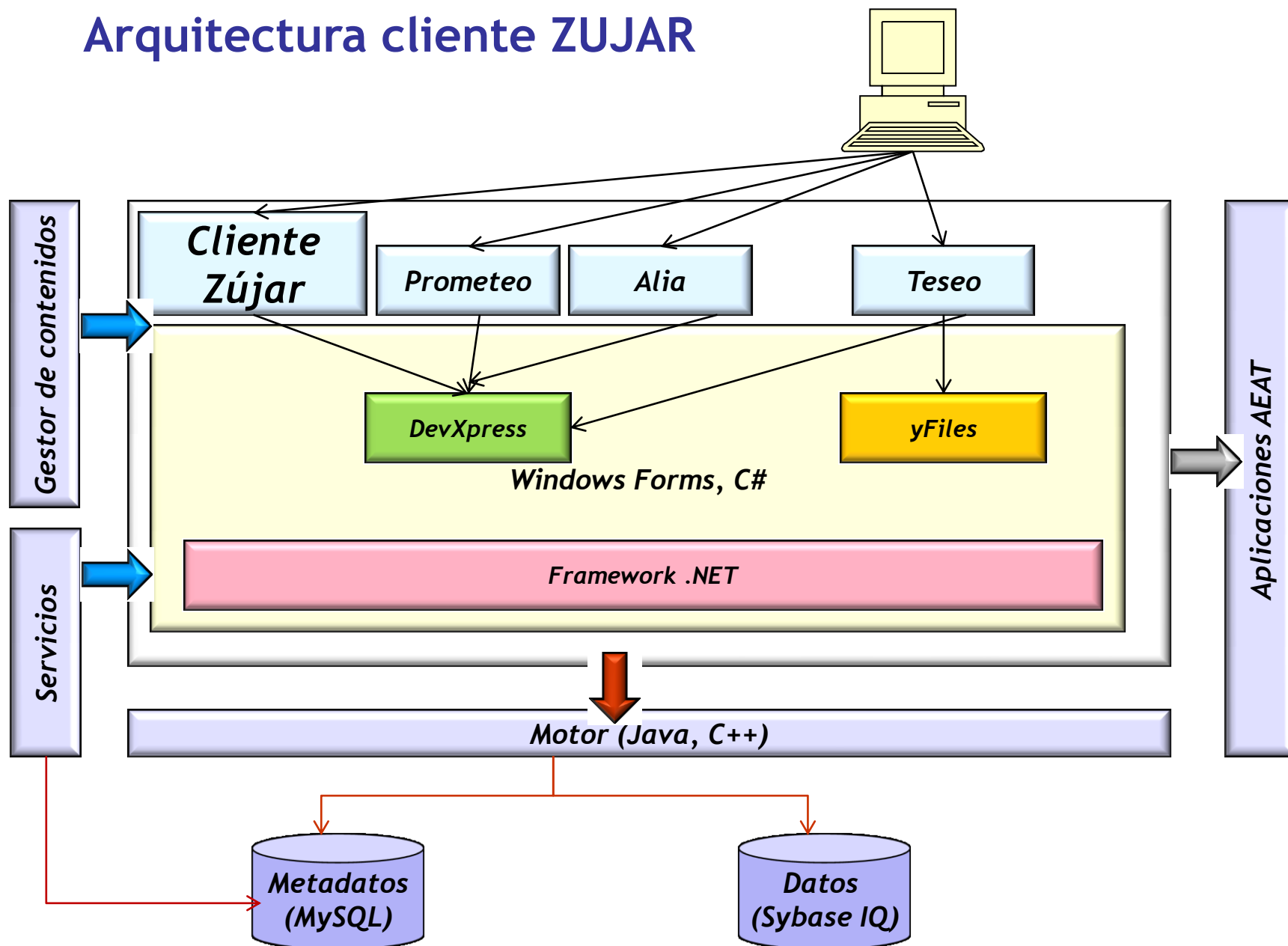
ZUAJR 2

¿Cuántos datos? Algunos ejemplos

- Tablas con más de 3.000 columnas
- Más de un millón de columnas definidas en más de 7.000 tablas.
- Más de 94.000.000.000 de filas
- Alrededor de 50 TB de disco
- Tablas físicas con más de 2.000.000.000 de filas
- Tablas lógicas divididas en tablas físicas y que en total contienen más de 19.000.000.000 de filas.
- Más de 500.000 consultas diarias, normalmente pesadas



Arquitectura cliente ZUJAR



Cliente Zújar

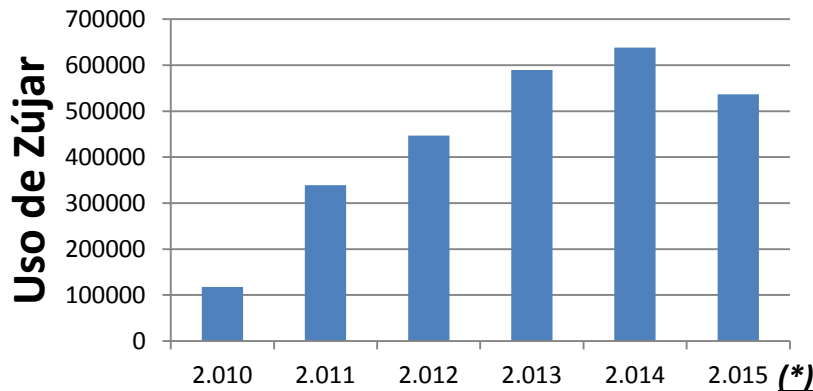
- Herramienta de desarrollo
 - Microsoft Visual Studio 2010
 - Frameworks 4.0
- Cliente pesado
 - Windows Forms
- Lenguaje de desarrollo
 - C#
- Componente de presentación de datos
 - Developer Express v2011
- Componente de visualización de grafos
 - yFiles .Net 4.2

Motor Zújar

- Desarrollo Java (frontal) y C++ en Linux (SUSE 11)
- Interfaz SOAP con aplicaciones clientes.
- “dialecto” SQL para dar respuesta a las necesidades específicas del negocio: *zSQL*.
- SGBD: MySQL y SAP Sybase IQ (futuro IMPALA).
- El Motor se encarga de transformar *zSQL* a SQL y resolver las peculiaridades de cada SGBD.

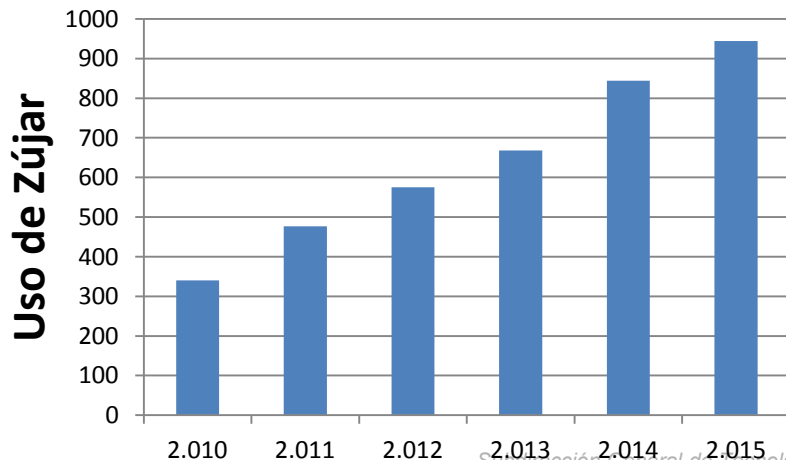


Número de peticiones del tipo "abrir consulta"

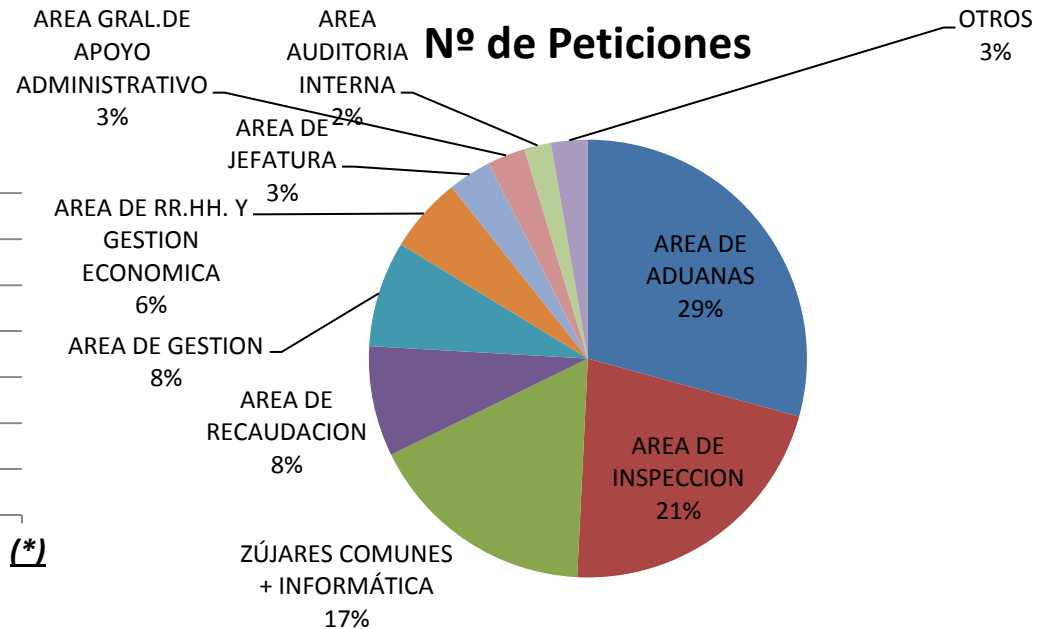


(*) 9 primeros meses

Número de Zújares distintos



Nº de Peticiones



Número de usuarios distintos

