

Universidad Nacional de Ingeniería
FACULTAD DE INGENIERIA INDUSTRIAL Y DE SISTEMAS



*Analizador de Datos Alfanuméricos
Aplicado a la Base de Datos Clientes
de un Banco*

INFORME DE INGENIERIA

Para optar el Título Profesional de:

INGENIERO DE SISTEMAS

Juan Carlos Camara Escudero

Lima - Perú
1996

ANALIZADOR DE DATOS ALFANUMERICOS
APLICADO A LA BASE DE DATOS CLIENTES
DE UN BANCO

DEDICATORIA

En nombre de mis padres, mis hijos
y esposa dedico el presente trabajo
a Dios.

Que me acompaña y me acompañará
siempre.

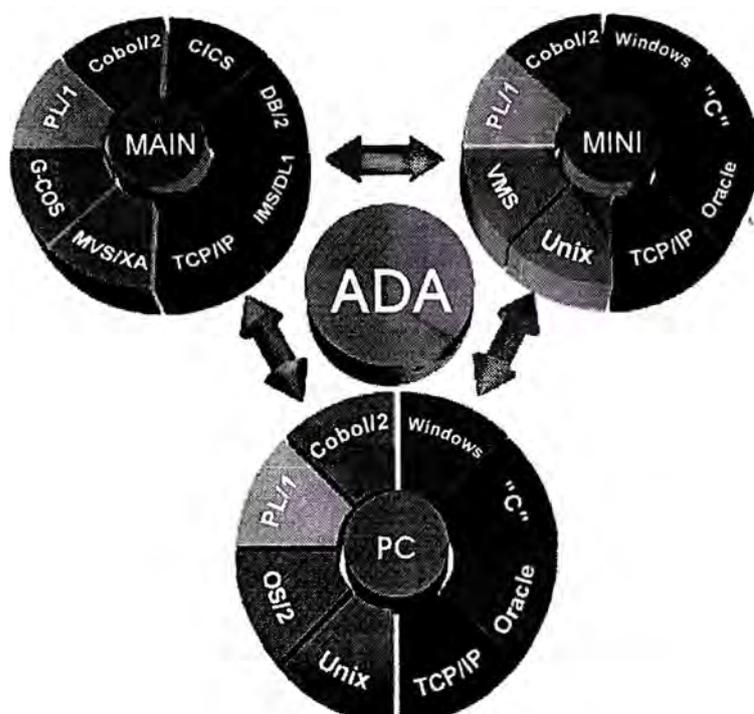
INDICE

SUMARIO.....	1
INTRODUCCION.....	3
1. PROBLEMAS.....	5
1.1 CAMPO DE ACCION.....	5
1.2 AMBITO DEL PROBLEMA.....	6
2. METODOLOGIA A EMPLEAR.....	9
2.1 MARCO CONCEPTUAL.....	9
2.2 ETAPAS O PASOS.....	11
3. DESARROLLO DE LA SOLUCION.....	18
4. RESULTADOS.....	36
5. ANALISIS COSTO-BENEFICIO.....	37
6. CONCLUSIONES Y RECOMENDACIONES.....	40
ANEXOS.....	42
BIBLIOGRAFIA.....	74

SUMARIO

El presente trabajo es un informe que resume los pasos a seguir en la normalización de archivos de una Base de Datos Clientes, en donde se encuentran errores de duplicidad de récord para el mismo cliente por errores de digitación en los datos personales del mismo. Esto se complica aun cuando el posible cliente, llamado también cliente potencial, efectúa operaciones como por ejemplo el envío de dinero a otros países sin estar todavía registrado en la Base de Datos, porque genera de las operaciones temporales con un gran porcentaje de error en lo que se refiere a la validez de los datos.

El proyecto comenzó en Setiembre de 1993 y con un tiempo previsto de desarrollo de 2 años, en el que se desarrolló un prototipo inicial y un sistema en lenguaje COBOL con DB2 para el ambiente 3090. Para las pruebas iniciales también fue creado un ambiente temporal (on Batch) en Foxpro para normalizar los datos de los clientes existentes. Esto fué hecho directamente en las oficinas de la Software House llevando los datos de los clientes del banco y retornándoles los mismos corregidos en parte con la intervención de los programas creados. Mi participación ha sido desde el comienzo del proyecto donde he tenido la posibilidad de partir del análisis hasta llegar al diseño y la respectiva programación.



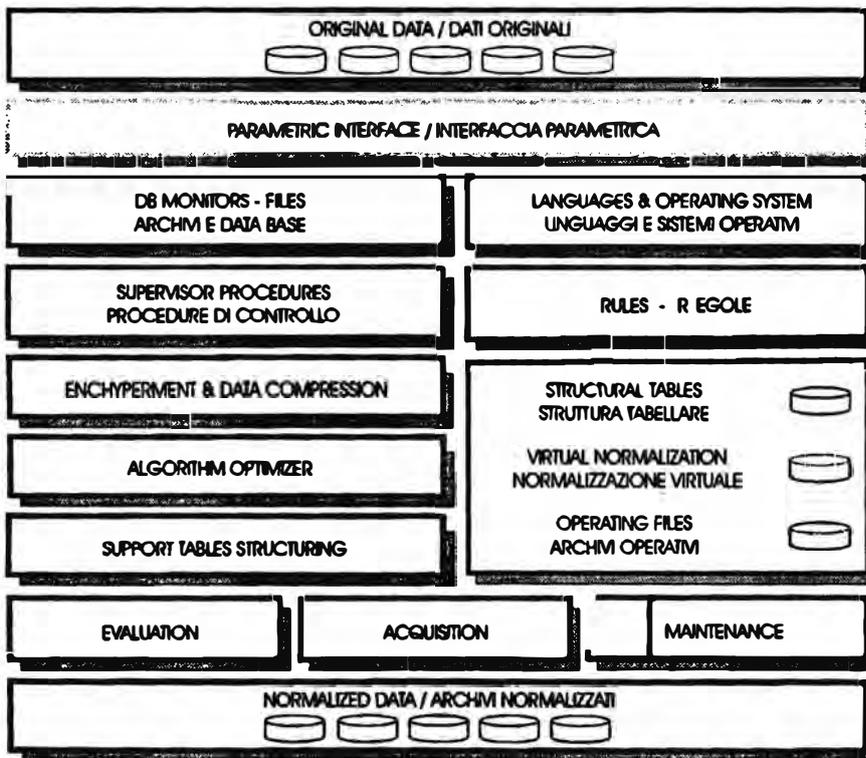
El banco en mención se llamaba antes "Cassa di Risparmio di Verona, Vicenza, Belluno y Ancona" y ahora "CariVerona Banca Spa" que pertenece al grupo UniBanca. La "CariVerona Banca Spa" está considerada actualmente como uno de los 4 primeros bancos de Italia y dentro de los 100 primeros bancos a nivel mundial.

Descriptores: Base de Datos, Banco, Algoritmo, Analizador de Datos.

INTRODUCCION

Uno de los aspectos más importantes de un Banco es el tratar de presentar una buena imagen hacia el cliente sea del punto de vista de seguridad económica como también de gestión al interno del mismo. Es decir, el Banco trata cada vez más de mejorar sus servicios, crear nuevos servicios alternativos que den la posibilidad a los clientes de efectuar el mayor número de operaciones en un menor tiempo, la posibilidad de realizarlas no solo a nivel nacional sino también en un area de mayor diámetro, al mismo tiempo dar la seguridad en cuanto a los intereses madurados. Estos objetivos que un Banco puede tener en el corto o largo plazo necesitan un gran soporte económico, es decir, que los costos operativos sean reducidos cada vez más dentro de sus posibilidades.

Reducir los costos internos operativos de un Banco puede significar un trabajo de optimización de las operaciones internas que se realizan y esto a su vez generar una mejor gestión de los datos que viajan en el sistema. Uno de los problemas que el Banco CariVerona Spa (Cassa di Risparmio di Verona) queria resolver en parte era la normalización de la Base de Datos Clientes, es decir, la eliminación de redundancias de registros que identifican al mismo cliente por efecto de errores de digitación en los datos personales tales como dirección, nombres y apellidos, localidad, y Código Postal (CAP). Estos errores se presentan debido al problema histórico de



Italia donde los nombres atribuidos a las calles después de la segunda guerra han sido dados en manera desordenada y donde la superpoblación existente ha llevado a complicar aún más las cosas.

Como referencia debemos tener presente que en Italia la posibilidad de cometer un error es grande porque existe una superpoblación. En términos numéricos, el área territorial italiana es de 301,225 Km² contra la peruana 1'285,215.6 Km². Además la densidad poblacional italiana es de aproximadamente 200 habitantes por cada Km² mientras la densidad de población peruana es de aproximadamente 20 habitantes por Km². Es decir, la cuarta parte del territorio peruano pero con aproximadamente 10 veces la densidad poblacional peruana (aproximadamente 3 veces la población peruana).

El análisis de Costo-Beneficio de implantar el sistema, da un resultado positivo. Si bien el costo inicial es alto, el beneficio será cada vez mayor cuando se incrementen el número de Clientes y sus movimientos. El presente sistema no requerirá entonces ningún cambio pero si aumentarán los beneficios para el banco.

1. PROBLEMAS

1.1 CAMPO DE ACCION

El campo de acción del sistema es tal de poder aplicar la solución a una serie de problemas y de exigencias típicas de la gestión de archivos del tipo **anagráfico**. Gracias a las técnicas de Software Engineering empleadas en el desarrollo de la aplicación, el sistema puede ser realizado in modo independiente de las características y del tipo de organización de los archivos tratados. El sistema, en su máxima expresión, presenta además la capacidad de consentir la creación automática de interrelaciones entre elementos específicos, de naturaleza y origen diferentes. Se trata de relaciones congruentes y lógicas que pueden ser aplicadas a elementos propios de operas literarias, de documentación técnica y manualística en general.

Por lo que se refiere a los aspectos aplicativos, en este banco, viene considerado principalmente todo lo que está relacionado con el tratamiento de los datos **anagráficos**. Al usuario viene pedida la eventual identificación y solución de aspectos diversos de los tratados, es decir, los diferentes elementos de aquellos típicos de los datos **anagráficos** Actividad que gracias a las características de flexibilidad y de parametrización del sistema, no presentan vínculos significativos. Las características de parametrización

le dan totalmente transparencia por lo que se refiere a su operatividad. Con esto se entiende el empleo de la mayor parte de las plataformas informáticas actualmente disponible en el mercado.

1.2 AMBITO DEL PROBLEMA

Como premisa se puede afirmar que, por su naturaleza, los registros civiles presentan el nivel más alto de redundancia porque generalmente son datos tratados con medios de elaboraciones electrónicas. Se trata de un tipo de información donde se pueden encontrar el mayor número de errores. Y es así que algunos datos como localidad, dirección, CAP y otros, vienen repetidos cada vez que se registran los datos civiles. Esto también es válido cuando estos datos son usados en otros registros. Ejemplo: S. Nazzaro dei Burgundi se repetirá, con las varias combinaciones, en todos los registros civiles relativos a los ciudadanos de tal localidad. Lo mismo sucede con la dirección, el CAP y otros elementos comunes codificados.

Se trata de una exigencia debida principalmente a la objetiva imposibilidad de atribuir un código de identificación a los elementos citados. Además de esto deben ser considerados vínculos de orden determinístico, exclusivamente a la búsqueda basada en métodos de acceso tradicionales. Condición que se convierte en una gran posibilidad de error y una menor

uniformidad (standarización) en el tratamiento de elementos comunes como por ejemplo : via, corso, avvocato, etc.

En los puntos sucesivos vienen indicados algunos elementos de base que contribuyen a conformar los registros civiles y hemos considerado ser los más significativos, los cuales derivan de tablas construídas automáticamente del sistema. Estos corresponden a los nombres, apellidos, títulos académicos, tipo y tantos otros atributos que contribuyen a formar un registro civil o datos personales de un individuo. Como es fácil intuir, se tratan de valores comprimidos, los cuales podrían generar valores de amplitud iguales a 4 o 6 posiciones numéricas.

Claramente el conocimiento de tal realidad es de poca ayuda si no se llega a atribuir automáticamente a cada elemento identificado un valor unívoco. Esto es debido principalmente al hecho que la escritura representa sonidos (fonemas) que tienen tiempos y valores similares .

A complicar las cosas, interviene el hecho que, también los fonemas, difícilmente pueden dar como resultado un valor cierto. Esto en particular cuando los elementos asociados (strings) son muchos (apellido, nombre, localidad, etc.). En el caso de los datos **anagráficos**, faltan también reglas fijas para definir la secuencia y las eventuales contradicciones de los varios componentes tratados. En realidad, un apellido y un nombre, en términos de

registración, pueden ser invertidos, precedidos, intercalados o seguidos de un título o un identificativo particular. Otra razón por la cual viene mayormente incontrolable el fenómeno es la presencia debida a las letras dobles, omitidas o introducidas erróneamente en el idioma italiano.

2. METODOLOGIA A EMPLEAR

2.1 MARCO CONCEPTUAL

La solución propuesta se basa en una serie de algoritmos y rigurosas reglas matemáticas de derivación de sistemas de *encryption* y de *data compressing*. El conocimiento de los elementos que conforman las localidades, direcciones, nombres y apellidos nos consiente de hipotetizar el campo de análisis capaz de atribuir el grado de certeza esperada. Esto representa la atribución automática de una pseudo codificación específica.

El campo de análisis, la frecuencia de los elementos y los métodos de registración de los datos forman las reglas mediante las cuales se determinan la auto-codificación de los elementos tratados. Esta considera las letras que componen un string las cuales vienen tratadas empleando en parte las fórmulas matemáticas propias de las más sofisticadas reglas de encriptación / decriptación. También vienen usadas técnicas de compresión de datos: como resultado final se obtiene el valor conseguido el cual asume características de univocidad y al mismo tiempo presenta grados de precisión tales de consentir el señalar errores de digitación o de transcripción.

Se trata de un proceso complejo, constituido de la capacidad de aislar elementos extraños, de excluir caracteres especiales y de eliminar espacios en blanco en exceso. Esto consiente de aislar in modo automático, los strings que constituyen, los elementos significativos de los datos **anagráficos**. Por lo que resta de cuanto expuesto podemos, a continuación indicar las áreas de aplicación de las actividades de normalización y de correlación de los archivos **anagráficos**.

La normalización de los archivos **anagráficos** permite de considerar un alto nivel cualitativo a los datos **anagráficos** tratados. Esto viene efectuado mediante la individualización y la consecuente corrección de los eventuales errores encontrados automáticamente del sistema. Se trata de una operación de vital importancia para la imagen de una institución. Además la corrección de los errores comporta un importante ahorro debido a documentos rechazados por el Correo o erróneamente entregados a causa de tales errores. En realidad, de la experiencia adquirida en el tratamiento de diferentes cantidades y tipologías de archivos, raramente hemos encontrado niveles de errores que resulten ser inferiores al 10% de la totalidad de los elementos tratados. Mediante el control de los errores evidenciados y la eventual corrección de los mismos, se obtiene un aumento considerable de la capacidad de correlación con archivos y datos **anagráficos**.

2.2 ETAPAS O PASOS

La capacidad de los algoritmos desarrollados y el empleo de las tablas de soporte, donde están los datos de atribución, las reglas de análisis y elaboración, permiten superar los vínculos de búsqueda de tipo determinístico, característico de los sistemas tradicionales. El resultado está representado por una autocodificación de los elementos **anagráficos** tratados, el cual grado de certeza es directamente proporcional al grado de correlación esperado. Las operaciones desarrolladas en esta fase tienden a la eliminación de posibles errores y de eventuales elementos de perturbación los cuales son

Caracteres especiales (puntos, comas, barras)

Atributos (Avvocato, Carrozeria, detto)

Espacios en blanco

Consonantes incongruentes

Una vez ejecutadas estas operaciones, los datos conseguirán una mayor congruencia y pueden ser empleados para la determinación de la clave de correlación. Esta viene determinada independientemente de la secuencia de registración de las palabras (el nombre puede preceder el apellido o viceversa) y del tamaño de las mismas (J.Carlos igual a Juan Carlos). A tal propósito sirve aclarar que, mediante la generación de la clave de

ORIGINAL FILES

FILES	CUSTOMER NAME	ADDRESS	ZIP	CITY
 <p>INTERNATIONAL SERVICES</p>	ARREND JOSEFINA SA GIUSEPPE BENTRALLI KLAUS SHUMAKER GMBH ROSSI MARIANO CON. ADELE BIANCHI	JARRON 27 VIA TURRINA, 15A VIENER STRABE 35/B P.BRA BURLINGTON ROAD	6012 67030 8024 35041 12020	TARRAGONA CANSANO VICENTINO KRUGERKIERKE BATTAGLIA WILLMINGTON MA
 <p>LOANS & DISCOUNTS</p>	ARREND JOSEFINA SA BENTRALLI GIUSEPPE FORNITURE WILLIAMSON CO KLAUS SHUMAKER GMBH MERCIER JEAN L ROSSI MARIANO	JARRON 27 V. TURRINA, N 15A BURLINGTON ROAD VIENER STRASSE 35/B 44 RUE DE LAGARTERRE	6012 67030 12020 80/24	TARRAGONA CANSANO V. WILLMINGTON MA KRUGKIERKE ST MALO ARR.9 BATTAGLIA T.
 <p>CONSUMER LOANS</p>	ARREND JOSEFINA SA BENTRALLI GIUSEPPE JEAN LUIS MERCIER KLAUS SHUMAKER GMBH ROSSI MARIANO	JARRON 27 VIA TURRINA, 15 44 RUE DE LA GARTERRE VIENER STRASSE 35/B P. BRA	6012 67030 80/24 35100 12020	TARRAGONA CANSANO SAINT MALO 9 KRUGERKIERKE BATTAGLIA T. WILLMINGTON MA
 <p>DEPOSIT SERVICES</p>	GIUSEPPE BENTRALLI JOAO FIGUREIDO JOSEFINA ARRENDORA KLAUS SHUMAKER GMBH ROSSI MARIANO	VIA TURRINA, 15A RUA BRANCA LUZ. ED.REAL CALLE JARRON 27 VIENER ST. 35/B PIAZZA BRA BURLINGTON ROAD	67030 6012 80/24 35041	CANSANO VICENTINO BELO HORIZONTE TARRAGONA KRUGERKIERKE BATTAGLIA T. WILLMINGTON MA

KEY ATTRIBUTION

CODE	CUSTOMER NAME	KEY	ADDRESS	KEY	ZIP	CITY	KEY
C01	ARREND. JOSEFINA SA	1327	CALLE JARRON N. 27	4344	6012	TARRAGONA	3391
C02	GIUSEPPE BENTRALLI	7732	VIA TURRINA, 15A	3331	67030	CANSANO VICENTINO	9812
C03	KLAUS SHUMAKER GMBH	1121	VIENER STRABE 35/8	2134	80/24	KRUGERKIERKE	3377
C04	ROSSI MARIANO CON. ADELE BIANCHI	3313	P. BRA	1221	35041	BATTAGLIA	3837
C05	WILLIAMSON FURNITURE CO.	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C11	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
C12	BENTRALLI GIUSEPPE	7732	V TURRINA, N. 15A	3331	67030	CANSANO V.	9812
C13	FORNITURE WILLIAMSON CO.	8812	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C14	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/8	2134	80/24	KRUGERKIERKE	3377
C15	MERCIER JEAN L.	8444	44 RUE DE LAGARTERRE	5641	ST MALO ARR. 9	1721
C16	ROSSI MARIANO	3313	P. BRA	1221	35041	BATTAGLIA	3837
C21	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
C22	BENTRALLI GIUSEPPE	7732	VIA TURRINA, 15	3331	67030	CANSANO	9812
C23	JEAN LUIS MERCIER	8444	44 RUE DE LA GARTERRE	5841	SAINT MALO	1721
C24	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/8	2134	80/24	KRUGERKIERKE	3377
C25	ROSSI MARIANO	3313	P. BRA	1221	35100	BATTAGLIA T.	3837
C26	WILLIAMSON FURNITURE	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C31	GIUSEPPE BENTRALLI	7732	VIA TURRINA, 15A	3331	67030	CANSANO VICENTINO	9812
C32	JOAO FIGUREIDO	9978	RUA BRANCA LUZ ED.REAL	7766	BELO HORIZONTE	2131
C33	JOSEFINA SA ARRENDORA	1327	CALLE JARRON 27	4344	6012	TARRAGONA	3391
C34	KLAUS SHUMAKER GMBH	1121	VIENER ST. 35/8	2134	80/24	KRUGERKIERKE	3377
C35	ROSSI MARIANO	3313	PIAZZA BRA	1221	35041	BATTAGLIA T.	3837
C36	WILLIAMSON FURNITURE CO.	6612	BURUTON ROAD	6541	12020	WILLMINGTON MA	6371

correlación, el archivo **anagráfico** adquiere una congruencia virtual que potencialmente consentiría de corregir eventuales errores y que, en cada caso, permitiría de correlacionar en modo automático y seguro, archivos derivados.

Las actividades de correlación, por sus características, vienen aplicadas entre archivos propios, vale decir entre archivos de aplicaciones diversas, o también entre archivos propios y archivos de derivación externa. En el primer caso se trata de actividades que generalmente tienden a uniformizar los archivos del banco y constituyen la fase que precede las actividades de normalización del archivo final uniformado. Se trata de actividades que consienten de generar una clave altamente confiable, la cual aumenta la posibilidad de correlación con archivos derivados.

En el segundo caso, los archivos externos, por su naturaleza y empleo, raramente pueden justificar la aplicación de operaciones de normalización. Esto es debido al hecho que la presencia de eventuales errores, señalados en automático por el sistema, no limita la capacidad de correlación. La clave de correlación en realidad, viene determinada en función del resultado esperado, donde en presencia de un hipotético nombre Cralos (errado) este será considerado como Carlos (correcto).

FIELD SELECTION

CODE	CUSTOMER NAME	KEY	ADDRESS	KEY	ZIP	CITY	KEY
003	KLAUS SHUMAKER GMBH	1121	VIENER STRABE 35/B	2134	80/24	KRUGERKIERKE	3377
014	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/B	2134	80/24	KRUGKIERKE	3377
024	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/B	2134	80/24	KRUGERKIERKE	3377
034	KLAUS SHUMAKER GMBH	1121	VIENER ST. 35/B	2134	80/24	KRUGERKIERKE	3377
001	ARREND. JOSEFINA SA	1327	CALLE JARRON N. 27	4344	6012	TARRAGONA	3391
011	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
021	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
033	JOSEFINA SA ARRENDORA	1327	CALLE JARRON 27	4344	6012	TARRAGONA	3391
004	ROSSI MARIANO CON. ADELE BIANCHI	3313	P. BRA	1221	35041	BATTAGLIA	3837
004	ADELE BIANCHI	3313	P. BRA	1221	35041	BATTAGLIA	3837
016	ROSSI MARIANO	3313	P. BRA	1221	35041	BATTAGLIA T.	3837
025	ROSSI MARIANO	3313	P. BRA	1221	35041	BATTAGLIA T.	3837
035	ROSSI MARIANO	3313	PIAZZA BRA	1221	35041	BATTAGLIA T.	3837
005	WILLIAMSON FURNITURE CO.	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
018	FURNITURE WILLIAMSON CO.	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
006	WILLIAMSON FURNITURE	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
036	WILLIAMSON FURNITURE CO.	6612	BURLITON ROAD	6541	12020	WILLMINGTON MA	6371
002	GIUSEPPE BENTRALLI	7732	VIA TURRINA, 15A	3331	67030	CANSANO VICENTINO	9812
012	BENTRALLI GIUSEPPE	7732	V TURRINA, N. 15A	3331	67030	CANSANO V	9812
022	BENTRALLI GIUSEPPE	7732	VIA TURRINA, 15	3331	67030	CANSANO	9812
016	MERCIER JEAN L.	3444	RUE DE LAGARTERRE	5641		ST MALO ARR 9	1721
023	JEAN LUIS MERCIER	3444	RUE DE LA GARTERRE	5641		SAINT MALO 9	1721

UNIFIED FILE

CODE	CUSTOMER NAME	KEY	ADDRESS	KEY	ZIP	CITY	KEY
C12,C22,C31	BENTRALLI GIUSEPPE	7732	VIA TURRINA N. 15A	3331	67030	CANSANO VICENTINO	9812
C32	JOAO FIGUEREIDO	9978	RUA BRANCA LUZ. ED. REAL	7766	BELO HORIZONTE	2131
C01,C11,C21,C33	JOSEFINA SA ARRENDORA	1327	CALLE JARRON 27	4344	6012	TARRAGONA	3391
C15, C23	JEAN LOUIS MERCIER	3444	44 RUE DE LG GARTERRE	6641	SAINT MALO ARROND. 9	1721
C04,0,16,C25,C35	ROSSI MARIANO	3313	PIAZZA BRA	1221	35041	BATTAGLIA TERME	3837
C03,C14,C24,C34	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/B	2134	80/24	KRUGERKTERKE	3377
C05,C13,C26,C36	WILLIAMSON FURNITURE CO.	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MASS	6371

DATA DICTIONARY

CODE	COSTUMER NAME
C12,C22,C31 C32 C01,C11,C21,C33 C15, C23 C04,0,16,C25,C35 C03,C14,C24,C34 C05,C13,C26,C36	C73.N39:T94,168(15A)L07 N12.C22:T83,166(ED. REAL)L21 N34.R11,31:T4,173(27)L10 C63.N70:T65,L54:L99 N88.N28,C23:(44)T41,142:L 11 N43,C51,R37:168,105(358)L84 C31.S45,R95:L92,T52:L49

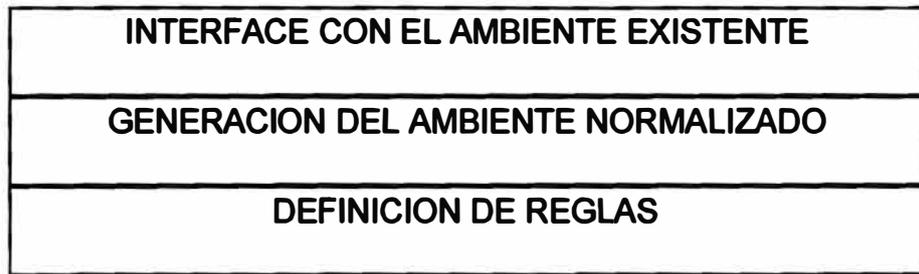
Las actividades de corrección de los errores evidenciados por el sistema, son realizados bajo el control del usuario, es decir con la autorización respectiva del usuario. En la parte relativa a esta actividad vienen suministradas las instrucciones operativas que han sido definidas para este fin. Estas actividades son del tipo opcional y pueden ser definidas en fase de instalación o cuando se retenga conveniente.

Además de las correcciones de los errores es posible standarizar los atributos por lo cual, una vez definidas las modalidades, palabras como *via*, *piazza*, *farmacia* y otras, pueden ser reportados en todos los datos **anagráficos** en el mismo modo. En este modo, *v.* puede ser siempre *via*, *p.zza* igual a *piazza* y *farm.* convertirse en *farmacia* o derepente mentenerse inalterada. Standarización valida también en sentido contrario, en este modo *via* puede convertirse en *v.*, *piazza* convertirse en *p.zza* y finalmente *farmacia* en *farm.*

Análogamente a la standarización de los atributos es posible realizar la contracción o la **ampliación** congruente de cualquier elemento tratado o serie de elementos cuales localidad o dirección. Otra ventaja está constituida en la posibilidad de ubicar los números cívicos en las posiciones consideradas convenientes y de agregar elementos que faltan como por ejemplo fracción u otros. Análogamente al aislamiento de los números

cívicos, es posible aislar y tratar oportunamente elementos referidos a condiciones como “coniugato”, “presso”, “in” y otros.

Tratándose principalmente de una metodología matemática, el sistema ha sido estructurado en función de tres componentes principales que a continuación vienen indicados:



Cada uno de los elementos identificados generalmente requieren, en fase de instalación, una serie de adaptaciones capaces de satisfacer las exigencias del usuario. En los puntos sucesivos vienen tratadas en forma detallada las operaciones que se realizaran, excluyendo todos aquellos pasos que se refieren físicamente a los aspectos del ambiente existente.

Interface con el Ambiente Existente

En lo concerniente con el ambiente existente, han sido aplicadas reglas de parametrización capaces de ser transparentes la adaptación del sistema al tipo de estructura y al dimensionamiento de los archivos existentes en el banco. En presencia de archivos manejados bajo el control de monitors DB del tipo relacional, el sistema viene suministrado con una interface predispuesta para tal fin. Esta interface viene generada según las reglas y los vínculos establecidos de parte del gestor del DB Empresarial.

Las funciones realizadas en el proceso de interface, se refieren a la gestión de los archivos en entrada y eventualmente a las actividades de restitución de la clave de relación. Las operaciones y las modalidades de restitución de las claves de relación vienen definidas en función de las reglas que conforman tal actividad.

En los casos en que los archivos vengán usados en forma diversa de los métodos relacionales, vienen establecidas las reglas para la gestión de los mismos según las exigencias del usuario. Generalmente, en estos casos, se opera con copias de los archivos originales en los cuales vienen realizadas las actividades de correlación y de eventualmente restitución de los archivos normalizados.

La planificación de las actividades de adaptación del sistema a las exigencias del usuario vienen realizadas en fase de instalación.

Normalización (navegación correlativa)

Los datos **anagráficos** contenidos en los archivos a tratar vienen reconocidos por sus características. Cada elemento aislado (objeto) viene explosionado encontrando lugar en posiciones específicas de memoria o recibiendo, por agregación, valores (pesos) cualificativos. Los objetos

explosionados pueden ser exactos o parcialmente errados. Tratándose de objetos que pueden asumir el valor de sinónimos.

El conjunto de elementos parcialmente completos, a través de una serie de programas de correlación lógica puede llevar a la corrección de estados formalmente de imposible solución según los métodos tradicionales de normalización. El método de solución viene denominado de navegación correlativa y en las tablas siguientes vienen indicadas las soluciones reales resueltas mediante el empleo de tal método.

3. DESARROLLO DE LA SOLUCION

Normalización (Definición)

Por normalización se debe entender todos las posibles intervenciones que permiten mejorar o corregir el contenido de archivos **anagráficos**. Las intervenciones puede ser de tipo automática o sino guiada. En los puntos sucesivos vienen dados algunos ejemplos de normalización, obtenidos automáticamente, como consecuencia de la aplicación de reglas predefinidas. Entonces la normalización se puede aplicar en las siguientes actividades

- Errores de digitación de los nominativos
- Errores de atribución CAP
- Direcciones erradas o incompletas
- Localidades erradas o incompletas
- Datos inexistentes
- Verificación del Código Fiscal
- Standarización del contenido de los campos anagráficos
- Correlación Alias/Sinónimos
- Solución de atribución de títulos y tipologías
- Standarización de atributos

Como se puede notar de la amplitud de los análisis posibles, se trata de una actividad aplicada principalmente a los archivos propios, en cuanto consiente de obtener una serie de ventajas las cuales son

- Eliminación datos redundantes
- Corrección de errores
- Evidenciación de las duplicidades
- Reorganización de los archivos
- Standarización datos y reglas
- Calidad de los archivos
- Congruencia de datos
- Correlación amplificada archivos externos

Las actividades de normalización pueden ser aplicadas a todos los elementos tratados con instrumentos de elaboración electrónica. En este caso vienen indicadas las reglas de normalización de tipo standard, que son disponible en la fase inicial.

En esta parte vienen presentados a manera de ejemplo, las reglas que comprometen los elementos de mayor importancia de un archivo **anagráfico** que son:

Razón Social,

Dirección,

CAP/Localidad.

Las reglas que vienen propuestas en estos ejemplos derivan de la aplicación del sistema a un gran número de archivos **anagráficos**. La significativa diversidad de organizaciones y de contenidos entre tales archivos, ha permitido de optimar los parámetros de análisis y de introducir reglas específicas. Con esta puntuación se trata de evidenciar que, a pesar de tratar con reglas optimadas, las exigencias y características de los archivos de los usuarios pueden sugerir la adopción de reglas y parámetros de análisis específicas.

Actividades Operativas

Aislamiento de errores ciertos: Se trata de los casos de ausencia de datos, de datos conflictivos o condiciones que presentan soluciones múltiples.

Identificación de las Reglas (Automatismos de norma): son las actividades aptas a identificar la distribución estadística de los varios "objetos" que conforman los elementos **anagráficos**. Actualmente han sido definidas en casi la totalidad de los casos y englobadas en las reglas de tratamiento de norma. Además de la distribución estadística, el sistema es en grado de señalar las posibles situaciones de conflicto, como el caso en el que **FR** normalmente interpretado como **frazione**, era la abreviación de **Francesco**.

Explosión de los Elementos: Las actividades de normalización se apoyan substancialmente en la capacidad de explosionar todos los componentes **anagráficos** y objetivizarlos con el acompañamiento de elementos de refuerzo particular. Estos son representados de exigencias de orden técnico como frecuencia, amplitud y colisión probabilística; y de características de tipo identificativo como toponimia, títulos académicos, **particelas** o similares. Tales elementos pueden ser ubicados en una posición específica del récord. Actividad que se convierte particularmente útil cuando se trata de **números cívicos**, condiciones de **presso** o tipo de clientela. La colocación de tales elementos, en la posición definitiva de los archivos, debe ser efectuada de acuerdo con las exigencias de parte del cliente. Estas se refieren asimismo a la standarización de toponimos, longitud de las descripciones de las localidades y de las direcciones y la ubicación de la eventual **frazione** (localidad).

Intervenciones correctivas : Se trata de todas las intervenciones por las cuales el sistema exige una solución externa (manual) en cuanto se llegan a encontrar soluciones múltiples. Esta intervención manual puede ser guiado de un programa. Es importante señalar que la solución de los errores no resolubles, relativamente a las direcciones, pueda contribuir a una reducción del número de errores de este tipo.

Autocodificación

Como consecuencia de las aplicaciones de las reglas y de los algoritmos de normalización o de correlación de los elementos **anagráficos** el sistema determina una pseudo autocodificación, la cual está dotada de un alto nivel de congruencia y de seguridad. En realidad las letras que constituyen los strings de datos, vienen tratados según las reglas de distribución estadística/matemática asociadas a una serie de números primos, mediante los cuales es posible determinar un resultado unívoco. Esto sucede independientemente de la secuencia en que los strings se presentan al interior de un determinado campo excluyendo o no la consideración de eventuales anagramas.

El resultado, de un récord específico generado en fase de análisis, constituye una clave que, en términos de valores relativos, presenta un nivel

de confianza que es directamente proporcional al grado de certeza esperada. Esta clave puede ser registrada directamente en el récord analizado o sino relacionado con tablas generadas con este propósito.

La construcción de la clave de relación se efectúa considerando separadamente los componentes homogéneos (campos) que constituyen el récord **anagráfico**. Tal separación viene además mantenida al interno de la clave de correlación con la cual resulta posible realizar eventuales comparaciones de relaciones referidas a uno o más campos de un mismo récord. En este ejemplo se ha tomado en consideración los campos que generalmente constituyen la parte más significativa de los archivos **anagráficos**.

Del análisis de la tabla siguiente se pueden ver las ventajas ofrecidas de la capacidad del sistema.

ARCHIVO ORIGINAL

	Razón Social	Dirección	Localidad
01	Mariano Rossani	Via Roma	Città di Cast.
02	Magazzini Bosco	Via Roma	C. di Castello
03	Rossani Mariano	Pza Bra	Battaglia T.
04	Bosco Mag.	Via Roma	Città di Castello
05	Mariano Rossani	Piazza Bra	Battaglia Terme
06	Rossani Mariano	Pza Bra	Battaglia T.
07	Magazzini Bosco	Borgo Real	Palo Vicentino
08	Mag. Bosco	Real Borgo	Palo Vic.
09	Rossani Mariano	Via Roma	Città Castello
10	Bosco Mag. Srl	V. Roma	Città di Castello

ARCHIVO NORMALIZADO

	Razón Social	Dirección	Localidad
01	Rossani Mariano	Via Roma	Città di Castello
09	Rossani Mariano	Via Roma	Città di Castello
03	Rossani Mariano	Piazza Bra	Battaglia Terme
05	Rossani Mariano	Piazza Bra	Battaglia Terme
06	Rossani Mariano	Piazza Bra	Battaglia Terme
02	Magazzini Bosco Srl	Via Roma	Città di Castello
04	Magazzini Bosco Srl	Via Roma	Città di Castello
10	Magazzini Bosco Srl	Via Roma	Città di Castello
07	Magazzini Bosco	Borgo Real	Palo Vicentino
08	Magazzini Bosco	Borgo Real	Palo Vicentino

Razón Social (Nominativos)

El término de Razón Social viene llamado a la parte de un archivo **anagráfico** que recoge los nominativos de Empresas o persona. En los sistemas de información estos elementos pueden ser indiferentemente el Cliente o el Deudor. Dadas las características de tales elementos y por la posibilidad ofrecida del sistema, las operaciones efectuadas para la creación de los strings son las siguientes

- Exclusión de los signos de puntuación
- Aislamiento de las iniciales
- Eliminación de Tipologías Empresariales
- Exclusión de atributos
- Exclusión de caracteres especiales
- Eliminación de espacios en blanco continuos

Los strings que constituyen los elementos del campo objeto de análisis, oportunamente elaborados, se presentan en las condiciones óptimas para la aplicación del algoritmo que genera la clave de relación. Además de la aplicación de las reglas, los strings son en este caso, adaptados a los eventuales standards de tratamiento definidos por parte del usuario. Con el término standard se debe entender en particular la guía de calles, los atributos y las Tipologías Empresariales eventuales.

La univocidad de los elementos que constituyen los datos **anagráficos** en su totalidad es directamente proporcional al número de elementos que la componen. A fin demostrativo vienen definidos algunos nominativos que serán analizados según las reglas establecidas.

En los ejemplos que a continuación se indica no viene relacionado el campo analizado (Razón Social) con los demás campos que por comodidad vienen indicados. Los strings aislados son aquellos señalados en **negrita**.

ARCHIVO ORIGINAL

	Razón Social	Dirección	Localidad
01	Mariano Rossani	Via Roma	Città di Cast.
02	Magazzini Bosco	Via Roma	C. di Castello
03	Rossani Mariano	Pza Bra	Battaglia T.
04	Bosco Mag.	Via Roma	Città di Castello
05	123Mariano Rossani	Piazza Bra	Battaglia Terme
06	Rossani Marianop	Pza Bra	Battaglia T.
07	Magazzini Bosco	Borgo Real	Palo Vicentino
08	Mag. Bosco F.Ili	Real Borgo	Palo Vic.
09	Rossani Mariano	Via Roma	Città Castello
10	Bosco Mag. Srl	V. Roma	Città di Castello

STRINGS

	Razón Social	Dirección	Localidad
01	Mariano Rossani	Via Roma	Città di Cast.
02	Bosco	Via Roma	C. di Castello
03	Rossani Mariano	Pza Bra	Battaglia T.
04	Bosco	Via Roma	Città di Castello
05	Mariano Rossani	Piazza Bra	Battaglia Terme
06	Rossani Mariano	Pza Bra	Battaglia T.
07	Bosco	Borgo Real	Palo Vicentino
08	Bosco	Real Borgo	Palo Vic.
09	Rossani Mariano	Via Roma	Città Castello
10	Bosco	V. Roma	Città di Castello

ARCHIVO NORMALIZADO

	Razón Social	Dirección	Localidad
01	Rossani Mariano	Via Roma	Città di Castello
09	Rossani Mariano	Via Roma	Città di Castello
03	Rossani Mariano	Piazza Bra	Battaglia Terme
05	Rossani Mariano	Piazza Bra	Battaglia Terme
06	Rossani Mariano	Piazza Bra	Battaglia Terme
02	Magazzini Bosco Srl	Via Roma	Città di Castello
04	Magazzini Bosco Srl	Via Roma	Città di Castello
10	Magazzini Bosco Srl	Via Roma	Città di Castello
07	Magazzini Bosco	Borgo Real	Palo Vicentino
08	Magazzini Bosco	Borgo Real	Palo Vicentino

Dirección

La dirección viene tratada, por la mayor parte de los casos, con las mismas reglas aplicadas con la Razón Social. Como es de notar, la denominación de las avenidas y calles, es decir la guía de calles, responde generalmente a requisitos de certeza y de univocidad al interno de una misma ciudad. Considerando estas premisas el sistema tendrá la capacidad de generar y de mantener todas las informaciones requeridas por la guía de calles, en modo congruente y con un nivel de seguridad alto.

La generación de la guía de calles resulta ser un subproducto de la aplicación del sistema y su empleo tiene la propiedad de otorgar un alto grado de seguridad de los datos tratados. En realidad, en el caso de eventuales errores de escritura, que vienen evidenciados automáticamente por el sistema, se puede proceder a la corrección directa, o sino guiada de los mismos.

Se trata de intervenciones de tipo relacional, entendiéndose con esto la eventual correlación de la guía de calles con el número de clientes que el usuario tiene en una dirección en particular. Tal capacidad no debe considerarse de tipo exclusiva. En realidad, dependiendo del número de clientes presentes en una determinada zona, resulta bastante difícil demostrar que existan, por hipótesis en una misma ciudad, dos avenidas

con nombres similares y que en una de estas los Clientes sean uno solo o pocos.

Las reglas para la determinación de los strings vienen aplicadas como indicadas en la parte relativa a la Razón Social excepto los casos de nombres conflictivos. Con este término se entienden casos como ***Via Vittorio Veneto, V.Vitt. Veneto, Via V. Veneto, Via Vittorio V.*** que deben ser tratados en forma específica. El conjunto de las operaciones aplicadas es el siguiente:

- Eliminación de los signos de puntuación
- Aislamiento de números romanos y arábigos
- Eliminación de las partículas
- Eliminación de las características
- Exclusión de las iniciales de los nombres
- Exclusión de los nombres conflictivos
- Exclusión de los caracteres no numéricos dentro del número cívico

CAP (Código di Avviamento Postale)

A efectos de análisis el CAP representa una ocasión para efectuar un eventual control formal en las registraciones y es de poca utilidad para la determinación de la clave de relación. En realidad, en muchos casos el CAP viene omitido en las registraciones y es atribuido o controlado por medio de archivos de soporte del tipo guía de calles o similares.

La verificación de la congruencia del CAP ayuda también a la posibilidad de aplicar controles formales de la Localidad y Dirección. Tal verificación se realiza recurriendo a la guía de calles la cual se ha hecho referencia anteriormente. Se trata de una posibilidad que debe ser considerada independientemente de las funciones de análisis aplicadas.

Las reglas que vienen aplicadas para la determinación de los strings relativos al CAP son :

- Eliminación de los signos de puntuación y caracteres especiales
- Eliminación de los espacios en blanco
- Análisis de la posición media

Cuando se encuentra la presencia de las condiciones indicadas en los puntos 1 y 2 el CAP es de poca validez. Para el punto 3 el control de la posición media se aplica, en ausencia de la guía de calles, para la capital y las ciudades de provincias. Ejemplo :

Venezia es la capital de la región del Veneto y tiene el CAP 30100, Verona como provincia de la misma región tiene el CAP 37100, otra provincia Bolzano tiene el CAP 39100.

Localidad

El campo relativo a la localidad viene tratado por la mayor parte de los casos con las mismas reglas de la dirección. Análogamente a la denominación de las avenidas y calles, la denominación de la localidad responde generalmente a requisitos de certeza y de univocidad. Se trata por esto de condiciones que permiten al sistema de generar y de mantener en el archivo atributos los nombres de todas las localidades que vienen analizadas.

La generación de las registraciones relativas a las Ciudades se convierte un subproducto de la aplicación del sistema y su uso tiene la propiedad de dar un alto grado de seguridad de los datos tratados. En realidad, en presencia de eventuales errores de escritura, que vienen evidenciados automáticamente por el sistema, se puede proceder a la corrección directa, o sino guiada de los mismos.

Se trata de intervenciones de tipo relacional, entendiéndose con esto la eventual correlación de las informaciones relativos a los nombres de las ciudades con el número de clientes que el usuario tiene en una ciudad en particular. Tal capacidad no debe considerarse de tipo exclusiva. En realidad, dependiendo del número de clientes presentes en una determinada ciudad, resulta bastante difícil demostrar que existan, por hipótesis en una misma ciudad, pocas unidades de Clientes.

Las reglas para la determinación de los strings vienen aplicadas como indicadas en la parte relativa a la Dirección excepto los casos de nombres conflictivos. Con este término se entienden casos como **Vittorio Veneto**, **Vitt. Veneto**, **V. Veneto**, **Vittorio V.** que deben ser tratados en forma específica. El conjunto de las operaciones aplicadas es el siguiente:

- Eliminación de los signos de puntuación
- Aislamiento de números romanos y arábigos
- Eliminación de las partículas
- Eliminación de las características
- Exclusión de las iniciales de los nombres
- Exclusión de los nombres conflictivos

Número Cívico

El número cívico en diversas aplicaciones resulta ser el elemento determinante por lo que se refiere a la generación de claves de relación.

Los casos de tratamiento pueden considerar la exclusión o no de tal elemento, independientemente del posicionamiento que este puede asumir al interno de un registro. El posicionamiento puede ser de tipo dedicado o relativo. En el primer caso se debe entender que ha sido definido un campo

para almacenar este elemento, en cambio el segundo es dividido con los demás elementos relativos a la dirección.

La función de aislamiento del número cívico viene efectuada según las reglas definidas para los otros elementos, vale decir que se excluyen los caracteres especiales, espacios en blanco y los signos de puntuación. La dimensión que normalmente viene considerada es de 4 posiciones con exclusiones de eventuales barras o números romanos, que en este caso pueden ser sustituidos de los equivalentes números arábigos.

Original Files

FILE	COSTUMER NAME	ADDRESS	ZIP	CITY
INTERNATIONAL SERVICES	ARREND. JOSEFINA SA	JARRON 27	6012	TARRAGONA
	GIUSEPPE BENTRALLI	VIA TURRINA, 15A	67030	CANSANO VICENTINO
	KLAUS SHUMAKER GMBH	VIENER STRABE 35/B	8024	KRUGERKIERKE
	ROSSI MARIANO CON ADELE BIANCHI	P.BRA	35041	BATTAGLIA
	WILLIAMSON FURNITURE CO.	BURLINGTON ROAD	12020	WILLMINGTON MA
LOANS	ARREND. JOSEFINA SA	JARRON 27	6012	TARRAGONA
	BENTRALLI GIUSEPPE	V TURRINA, N. 15A	67030	CANSANO V.
	FURNITURE WILLIAMSON CO.	BURLINGTON ROAD	12020	WILLMINGTON MA
	KLAUS SHUMAKER GMBH	VIENER STRASSE 35/B	8024	KRUGKIERKE
	MERCER JEAN L.	44 RUE DE LA GARTERRE		ST MALO ARR.B
ROSSI MARIANO	P.BRA	35041	BATTAGLIA T.	
CONSUMERS	ARREND. JOSEFINA SA	JARRON 27	6012	TARRAGONA
	BENTRALLI GIUSEPPE	VIA TURRINA, 15	67030	CANSANO
	JEAN LOUIS MERCER	44 RUE DE LA GARTERRE		SAINT MALO 9
	KLAUS SHUMAKER GMBH	VIENER STRASSE 35/B	8024	KRUGERKIERKE
	ROSSI MARIANO	P.BRA	35100	BATTAGLIA T.
WILLIAMSON FURNITURE	BURLINGTON ROAD	12020	WILLMINGTON MA	
DEPOSIT SERVICES	GIUSEPPE BENTRALLI	VIA TURRINA, 15A	67030	CANSANO VICENTINO
	JOAO FIGUREIDO	RUA BRANCA LUZ ED.REAL		BELO HORIZONTE
	JOSEFINA SA ARRENDORA	CALLE JARRON 27	6012	TARRAGONA
	KLAUS SHUMAKER GMBH	VIENER ST. 35/B	8024	KRUGERKIERKE
	ROSSI MARIANO	PIAZZA BRA	35041	BATTAGLIA T.
WILLIAMSON FURNITURE CO.	BURLINGTON ROAD	12020	WILLMINGTON MA	

Key Attribution

C01	ARREND. JOSEFINA SA	1327	CALLE JARRON N. 27	4344	6012	TARRAGONA	3391
C02	GIUSEPPE BENTRALLI	7732	VIA TURRINA, 15A	3331	67030	CANSANO VICENTINO	9812
C03	KLAUS SHUMAKER GMBH	1121	VIENER STRABE 35/B	2134	B0/24	KRUGERKIERKE	3377
C04	ROSSI MARIANO CON. ADELE BIANCHI	3313	P.BRA	1221	35041	BATTAGLIA	3837
C05	WILLIAMSON FURNITURE CO.	6812	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C11	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
C12	BENTRALLI GIUSEPPE	7732	V TURRINA, N. 15A	3331	67030	CANSANO V.	9812
C13	FURNITURE WILLIAMSON CO.	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C14	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/B	2134	B0/24	KRUGKIERKE	3377
C15	MERCIER JEAN L.	8444	44 RUE DE LAGARTERRE	5641		ST MALO ARR.9	1721
C16	ROSSI MARIANO	3313	P.BRA	1221	35041	BATTAGLIA T.	3837
C21	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
C22	BETRALLI GIUSEPPE	7732	VIA TURRINA, 15	3331	67030	CANSANO	9812
C23	JEAN LOUIS MERCIER	8444	44 RUE DE LA GARTERRE	5641		SAINT MALO 9	1721
C24	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/B	2134	B0/24	KRUGERKIERKE	3377
C25	ROSSI MARIANO	3313	P.BRA	1221	35100	BATTAGLIA T.	3837
C26	WILLIAMSON FURNITURE	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C31	GIUSEPPE BENTRALLI	7732	VIA TURRINA, 15A	3331	67030	CANSANO VICENTINO	9812
C32	JOAO FIGUREIDO	9978	RUA BRANCA LUZ ED.REAL	7768		BELO HORIZONTE	2131
C33	JOSEFINA SA ARRENDORA	1327	CALLE JARRON 27	4344	6012	TARRAGONA	3391
C34	KLAUS SHUMAKER GMBH	1121	VIENER ST. 35/B	2134	B0/24	KRUGERKIERKE	3377
C35	ROSSI MARIANO	3313	PIAZZA BRA	1221	35041	BATTAGLIA T.	3837
C36	WILLIAMSON FURNITURE CO.	6612	BURLITON ROAD	6541	12020	WILLMINGTON MA	6371

Field Selection

C03	KLAUS SHUMAKER GMBH	1121	VIENER STRABE 35/B	2134	B0/24	KRUGERKIERKE	3377
C14	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/B	2134	B0/24	KRUGKIERKE	3377
C24	KLAUS SHUMAKER GMBH	1121	VIENER STRASSE 35/B	2134	B0/24	KRUGERKIERKE	3377
C34	KLAUS SHUMAKER GMBH	1121	VIENER ST. 35/B	2134	B0/24	KRUGERKIERKE	3377
C01	ARREND. JOSEFINA SA	1327	CALLE JARRON N. 27	4344	6012	TARRAGONA	3391
C11	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
C21	ARREND. JOSEFINA SA	1327	JARRON 27	4344	6012	TARRAGONA	3391
C33	JOSEFINA SA ARRENDORA	1327	CALLE JARRON 27	4344	6012	TARRAGONA	3391
C04	ROSSI MARIANO CON. ADELE BIANCHI	3313	P.BRA	1221	35041	BATTAGLIA	3837
C04	ADELE BIANCHI	3313	P.BRA	1221	35041	BATTAGLIA	
C16	ROSSI MARIANO	3313	P.BRA	1221	35041	BATTAGLIA T.	3837
C25	ROSSI MARIANO	3313	P.BRA	1221	35100	BATTAGLIA T.	3837
C35	ROSSI MARIANO	3313	PIAZZA BRA	1221	35041	BATTAGLIA T.	3837
C05	WILLIAMSON FURNITURE CO.	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C13	FORNITURE WILLIAMSON CO.	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C26	WILLIAMSON FURNITURE	6612	BURLINGTON ROAD	6541	12020	WILLMINGTON MA	6371
C36	WILLIAMSON FURNITURE CO.	6612	BURLITON ROAD	6541	12020	WILLMINGTON MA	6371
C02	GIUSEPPE BENTRALLI	7732	VIA TURRINA, 15A	3331	67030	CANSANO VICENTINO	9812
C12	BENTRALLI GIUSEPPE	7732	V TURRINA, N. 15A	3331	67030	CANSANO V.	9812
C22	BETRALLI GIUSEPPE	7732	VIA TURRINA, 15	3331	67030	CANSANO	9812
C15	MERCIER JEAN L.	8444	44 RUE DE LAGARTERRE	5641		ST MALO ARR.9	1721
C23	JEAN LOUIS MERCIER	8444	44 RUE DE LA GARTERRE	5641		SANT MALO.9	1721

Microsoft FoxPro

Filter Ricerca Edit

Anagrafica		Selezione					
nome normalizzato		Anno	Contrib.	Nome normalizzato	Cap	Indirizzo	
FALCO DI GIANNI EMANUELE			00304141	FALCO DI GIANNI EMANUELE	00000		
FALCO DI GIANNI EMANUELE			00304041	FALCO DI GIANNI EMANUELE	00000		
FALLACI ADA							
FALLACI BRUNO							
FALLONI ALESSANDRO							
FALORSI ANTONIO							
FALORSI ANTONIO							
FALOSI VESSERO							

Main Record

Anno documento: Flag: Entrate patrimonio

Contribuente: 00304141

Nome: FALCO DI GIANNI EMANUELE

Estensione:

Data Nascita: 06041966

Luogo Nascita: MONTALCINO

Codice Fiscale:

Tipo Societa': 00

Tipo Soggetto: 1

Sesso: M stesso record

Codice Provincia: 000

Codice Comune: 000

Cap: 00000

Localita: MONTALCINO

Toponimo: VIA

Indirizzo: GIUSEPPE VERDI

Numero Civico: 17

Nascita (da nom.):

Detto:

Presso: 1

Frazione (da ind.):

Selezione

Anno documento: Flag: Entrate patrimonio

Contribuente: 00304141

Nome: FALCO DI GIANNI EMANUELE

Estensione:

Data Nascita: 06041966

Luogo Nascita:

Codice Fiscale:

Tipo Societa': 00

Tipo Soggetto: 1 stesso record

Sesso:

Codice Provincia: 000

Codice Comune: 000

Cap: 00000

Localita:

Toponimo:

Indirizzo:

Numero Civico:

Nascita (da nom.):

Detto:

Presso:

Frazione (da ind.):

Main

Clear

Move

Ok

Delete

Undelete

Fine

VISTA LOGICA

Filter Ricerca Edit

Anagrafica		Selezione					
Nome normalizzato	D	Anno	Contrib.	Nome normalizzato	Cap.	Indirizzo	D
AZZURRINI TERZILIO			00003140	BAGLIONI ILIO	00000		
AZZURRINI ZAJRA			00250251	BAGLIONI ILIO		VIA RICASOLI 12,	C
B O RTINI CELSO			00250001	BAGLIONI ILIO	00000		
BACCI FABIO			00250183	BAGLIONI MARGHERITA	00000		
BACHI PIERO			00002451	BAGLIONI PIERINA IN RENAI	00000		
BADI SABATINO			00250182	BAGLIONI MARGHERITA VED BALDI	00000		
BAGLIONI ARMIDO							
BAGLIONI ILIO							

Main Record		Selezione	
Anno documento:	Flag: Entrate patrimonio	Anno documento:	Flag: Entrate patrimonio
Contribuente:	00250251	Contribuente:	00003140
Nome:	BAGLIONI ILIO	Nome:	BAGLIONI ILIO
Estensione:		Estensione:	
Data Nascita:	29061930	<input checked="" type="checkbox"/> Data Nascita:	29061930
Luogo Nascita:	GAIOLE	<input checked="" type="checkbox"/> Luogo Nascita:	GAIOLE
Codice Fiscale:		Codice Fiscale:	
Tipo Societa':	00	Tipo Societa':	00
Tipo Soggetto:	1	Tipo Soggetto:	1
Sesso:		Sesso:	
Codice Provincia:	075	Codice Provincia:	000
Codice Comune:	540	Codice Comune:	000
Cap:	53019	Cap:	00000
Localita':	CASTELNUOVO BERARDENGA	Localita':	
Toponimo:	VIA	Toponimo:	
Indirizzo:	RICASOLI	Indirizzo:	
Numero Civico:	12	Numero Civico:	
Nascita (da nom.):		Nascita (da nom.):	
Detto:		Detto:	
Presso:		Presso:	
Frazione (da ind.):		Frazione (da ind.):	

VISTA LOGICA

Filter Ricerca Edit

Anagrafica

Nome_new	Indi	Loca	Cap_n	Via_n	Indi_new	Nro_n	Loca_new	FI
BABAZZI ROMANO	FRENZE VLE D AMICIS 39			VLE	D AMICIS	39		N
BABBONI ADORNA	VIA AFRICO 2 PRATACCIO			VIA	AFRICO	2		N
BABETTO GASTONE	V.BARETTI 9 PADOVA		35100	VIA	BARETTI	9	PADOVA	X
BABETTO GUIDO	VIA MONTE GALLO 27 PADOVA			VIA	MONTE GALLO	27		N
BABETTO MARINA	V.RICATI 25 PADOVA			VIA	RICATI	25		N
BABETTO ROBERTO	V.CIAMCIAN 80 PADOVA			VIA	CIAMCIAN	80		N
BABETTO ROBERTO	V.PIOVEGHETTO S.N.C. PADOVA			VIA	PIOVEGHETTO C. PAIS.N.			N
BABETTO ROBERTO	V.PIOVEGHETTO 28 PADOVA		35136	VIA	PIOVEGHETTO	28	PADOVA	U
BABIC MARTINO	VIA CANALETTO 1			VIA	CANALETTO	1		N

Selezione

Nome_new	Cap	Via	Indi_new	Nro	Loca_new
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	29	PADOVA
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA

Main Record

Nome_new: BABETTO ROBERTO
 Indi: V.PIOVEGHETTO 28 PADOVA
 Localita:
 Cap_new: 35136
 Via_new: VIA
 Indi_new: PIOVEGHETTO
 Num_new: 28
 Loca_new: PADOVA

Selezione

Nome_new: BABETTO ROBERTO
 Cap_new: 35136
 Via_new: VIA
 Indi_new: PIOVEGHETTO
 Num_new: 29
 Loca_new: PADOVA

Move Save Find NextF Fine

VISTA LOGICA

Filter Ricerca Edit

Anagrup						
Nome_new	Indi	Loca	Cap_n	Via_n	Indi_new	Nro_n
BABETTO ROBERTO	V.CIAMICIAN 80 PADOVA			VIA	CIAMICIAN	80
BABETTO ROBERTO	V.PIOVEGHETTO S.N.C. PADOV			VIA	PIOVEGHETTO C. PADOV	S.N.
BABETTO ROBERTO	V.PIOVEGHETTO 28 PADOVA		35136	VIA	PIOVEGHETTO	28 PADOVA

Selezione						
Nome_new	Cap	Via	Indi_new	Nro	Loca_new	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	29	PADOVA	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA	
BABETTO ROBERTO	35136	VIA	PIOVEGHETTO	28	PADOVA	

Main Record	Selezione
Nome_new: BABETTO ROBERTO	Nome_new: BABETTO ROBERTO
Indi: V.CIAMICIAN 80 PADOVA	
Localita: [REDACTED]	
Cap_new: [REDACTED]	<input checked="" type="checkbox"/> Cap_new: 35136
Via_new: VIA	<input checked="" type="checkbox"/> Via_new: VIA
Indi_new: CIAMICIAN	<input checked="" type="checkbox"/> Indi_new: PIOVEGHETTO
Num_new: 80	<input checked="" type="checkbox"/> Num_new: 28
Loca_new: [REDACTED]	<input checked="" type="checkbox"/> Loca_new: PADOVA

Move Save Find NextF Fine

Filter Ricerca Edit

Anagrup

Nome_new	Indi	Loce	Cap_n	Via_n	Indi_new	Nro_n	Loce_new	FI
RABAGLIETTI RODOLFO	VIA CASTELBARCO 3 MILAN			VIA	CASTELBARCO	3		N
RABAGLIETTI RODOLFO	V.CASTELBARCO 3 MILANO			VIA	CASTELBARCO	3		N

Localita			Strade		
Cap	Localita	Comune	Cap	Via	Indirizzo
20060	MIGNETE	ZELO BUON PER	20136	VIA	BOSELLI ANTONIO
15100	MILANA	ALESSANDRIA	20136	VLE	BRAHMS GIOVANNI
10040	MILANERE	ALMESE	20136	VIA	BRIOSCHI FRANCESC
89050	MILANESI	CALANNA	20136	VIA	BRUNACCI VINCENZO
10084	MILANI	FORNO CANAVES	20136	VIA	CAIMI GIUSEPPE
20095	MILANINO	CUSANO MILANIN	20136	PZA	CARO TITO LUCREZIO
20100	MILANO		20136	VIA	CASTELBARCO G.
20090	MILANO DUE	SEGRATE	20136	VIA	CASTIGLIONI LUIGI
48016	MILANO MARITTIMA	CERVIA MILANO M	20136	VIA	COL BRICON
20090	MILANO SAN FELICE	SEGRATE	20136	VIA	COL DEL ROSSO
20080	MILANO TRE		20136	VLE	COL DI LANA
20090	MILANO FIORI	ASSAGO	20136	VIA	COL MOSCHIN

Main Record		Selezione	
Nome_new:	RABAGLIETTI RODOLFO	Nome_new:	
Indi:	VIA CASTELBARCO 3 MILANO		
Localita:			
Cap_new:		<input checked="" type="checkbox"/> Cap_new:	
Via_new:	VIA	<input checked="" type="checkbox"/> Via_new:	
Indi_new:	CASTELBARCO	<input checked="" type="checkbox"/> Indi_new:	
Num_new:	3	<input checked="" type="checkbox"/> Num_new:	
Loce_new:		<input checked="" type="checkbox"/> Loce_new:	

4. RESULTADOS

Al implementar el Analizador de Datos Alfanuméricos se satisfizo los requerimientos de la Gerencia de Marketing, de llevar la administración de los datos personales de los Clientes y determinar el volumen de clientes potenciales (posibles clientes) que hacen operaciones en el banco pero que todavía no se deciden a establecer un contacto formal o continuativo con el banco. Es decir mediante este sistema el banco puede captar un mayor número de clientes y crear una buena imagen hacia los mismos.

Además los sectoristas también pueden decidir mas rápido sobre las autorizaciones o rechazos de las operaciones y también efectuar un control efectivo sobre las exigencias del banco con los organismos de control como por ejemplo la Superintendencia de Banca y Seguros, Ministerio de Finanzas, sobre las informaciones del tipo anti-narcotrafico o anti-reciclaje de dinero ya que el cliente es identificado en manera unívoca.

5. ANALISIS COSTO-BENEFICIO

Mediante este análisis se justificará la inversión de desarrollar el Analizador de Datos Alfanuméricos, los costos y tiempos que implican, y los beneficios que se obtienen.

Los costos del proyecto dependen de si se desarrolla por el propio personal de la empresa o si se encarga a terceros. En este caso se desarrolló el sistema por el personal externo.

Existen problemas con determinar los costos exactos en que se incurrieron porque por ejemplo el personal encargado no lo hizo en tiempo exclusivo, el computador para la implementación ya existía y también estaba dedicado a otros aplicativos, se usaron las instalaciones del banco y otros recursos que no fueron para el desarrollo del Analizador de Datos Alfanuméricos.

Para tal efecto se han estimado los tiempos y costos promedios para cada etapa del desarrollo del sistema y se han evaluado de acuerdo a los costos promedio del promedio del mercado.

Dentro de los recursos que se emplearon, las horas/hombre son las que se determinaron con mayor facilidad. El cuadro siguiente muestra las personas involucradas y su costo.

COSTOS DEL SISTEMA

ETAPA	DURACION
Análisis	8 semanas
Diseño	9 semanas
Implementación	10 semanas

PERSONA	COSTO/SEMANAL
Jefe de Proyectos	\$ 800
Analista de Sistemas	\$ 600
Analista Programador	\$ 400

ETAPA	PERSONAS	HORAS/ DIA	DURACION (semanas)	DURACION NETA (semanas)	COSTOS
Análisis	1 Jefe de Proyectos	4	8	4	\$ 3,200
	1 Analista de Sistemas	8		8	\$ 4,800
Diseño	1 Jefe de Proyectos	4	9	4.5	\$ 3,600
	1 Analista de Sistemas	8		9	\$ 5,400
Implementación	2 Analista de Sistemas	4	10	5	\$ 6,000
	4 Analistas Programadores	8		10	\$ 16,000

TOTAL : \$ 42,000

Los beneficios que se obtienen con el Analizador de Datos Alfanuméricos son diversos aunque difíciles de cuantificarlos en vista de que se trata de un sistema de información, y como tal proporciona información útil para la toma de decisiones, esta información tiene un alto valor pero difícil de cuantificar porque depende de varios factores como del momento en que se entregue, de su calidad y de consistencia.

La administración de los datos personales de los clientes resulta más sencilla puesto que se dispone de una información normalizada, es decir con datos standards que ayudan a evitar datos parciales que puedan confundir la procedencia del cliente.

Todos estos beneficios justifican los costos del desarrollo del sistema y a su vez que es necesario.

Lo principal es que el sistema proporciona la información inmediata para los distintos niveles del banco: operacional, táctico y estratégico.

6. CONCLUSIONES Y RECOMENDACIONES

El sistema diseñado e implementado satisface los requerimientos del Banco de procesamiento de datos en línea, información actualizada constante, identificación univoca del cliente, soporte para la toma de decisiones de los administradores, prestar un mejor servicio a los clientes y suministrar información a los organismos que controlan la banca.

La metodología empleada para el desarrollo del sistema sugiere el uso de un software (Case) para cada etapa de desarrollo del sistema (Análisis, Diseño e Implementación) con el fin de reducir tiempos en cada etapa, hacerlos mas consistentes y minimizar errores. En el caso del banco no se empleó esta herramienta en vista de que en el sistema predomina la parte matemática y los tiempos de respuesta pueden incrementarse. Lo recomendable para estos sistemas es que se desarrollen en lenguajes que generen códigos ejecutables pequeños para obtener mayor velocidad.

Como la implementación del sistema es independiente del análisis y diseño realizado, puede construirse en cualquier arquitectura computacional y con cualquier software de desarrollo de sistemas; el análisis y diseño es valido para cualquier banco y puede implementarse en cualquier computador, pero debe tener alta velocidad y capacidad en vista de la magnitud de clientes que pueden existir y el tipo de transacciones que se realizan.

En el caso del banco que se implementó, las transacciones se procesan en forma descentralizada en el computador ubicado en el Centro de Cómputo

del Banco con redes locales instaladas en cada agencia que se encarguen de procesar las transacciones y solo se accede al computador central para los accesos a las Bases de Datos de esta forma el proceso será mas rápido en las agencias y el computador central quedará menos congestionado.

ANEXOS

Anagráfico

Se entiende por anagráfico a los datos personales de un cliente o posible cliente tales como nombres y apellidos, dirección, localidad, Código Postal, etc. También viene interpretado como Base de Datos Clientes.

CAP (Códice di Avviamento Postale)

Viene interpretado como código postal donde cada zona geográfica tiene uno y solo un número definido.

Código de agrupación Bancaria en función de grupos económicos que dominan el mercado.

Son las iniciales de Asociación de Bancos Italianos

Particelas

Particelas debe ser interpretado como las partes acompañantes de un string, por ejemplo en el caso de un nombre "Maurizio di Lauro", la palabra "di" es una particella. En el caso de "F.Ili Nicolini" la palabra "F.Ili" es una particella.

Presso

Se denomina presso a la persona propietaria del inmueble o en el caso de departamento alquilado la persona que ha firmado el contrato, es decir, si se manda una carta a un cliente que ocupa un departamento que no es el suyo tiene que indicarse el nombre de la persona que tiene una relación directa con el inmueble (propietario o inquilino).

Número Cívico

Se denomina número cívico a la parte numérica de la dirección del cliente.

Campo

Como terminología nos referimos a las reglas informáticas, las cuales lo definen como una porción de registración efectuada en un dispositivo

magnético. Tal porción representa una parte específica de una serie de otros elementos (campos) relacionados entre sí. El agrupamiento de estos campos a veces constituyen lo que viene llamado Record o Registro. Las características de los campos vienen definidas en función de las exigencias de la aplicación utilizada. A continuación una breve descripción de las mismas:

Longitud	Se refiere al número máximo de posiciones asignadas
Numérico	Indica si el campo está constituido solo de números
Alfanumérico	Indica si el campo esta constituido de números y letras
Fecha	Es la fecha de calendario, expresada generalmente en día-mes-año
Nombre	Es el nombre lógico atribuido al campo
Availability	Indica si se trata de registración obligatoria

RECORD LOGICO

Codigo	Razón Social	Dirección	Localidad	CAP
---------------	---------------------	------------------	------------------	------------

Record o Registro

Por Record o registro, como indicado en el punto anterior, se entiende un grupo de campos que tienen un cierto parentesco entre ellos. Generalmente estos elementos representan una registración de tipo contable o de tipo gestional, como por ejemplo, cobros o movimientos de almacén. En este caso, por la amplitud del argumento, no se hará referencia a sistemas de tratamiento de datos, realizados en ambiente de tipo DB relacionales aunque si esta metodología es ampliamente empleada en módulos específicos del sistema.

El tipo de metodología aplicada para el tratamiento de los archivos electrónicos no representa un límite para el sistema. En fase de instalación vienen requeridas las características de los archivos a tratar en entrada y vienen creadas las condiciones para transportar tales archivos en ambiente original o sino en un nuevo ambiente. Esto naturalmente con el respeto de las reglas y standard propios de la plataforma operativa.

TIPOS DE RECORDS

Codigo	Razón Social	Dirección	Localidad	CAP
---------------	---------------------	------------------	------------------	------------

Num.	Nombres	Apellidos	Importe	Vencimiento
-------------	----------------	------------------	----------------	--------------------

Data Base

Se trata de un conjunto de informaciones manejados a través de los sistemas informativos. Estas vienen expuestas según las exigencias propias de las aplicaciones desarrolladas y responden a requisitos informáticos precisos.

La gestión de los archivos, según las técnicas de Data Base, vienen generalmente administradas por sistemas que se llaman Monitor DB. En muchos casos, en realidad, existe una interpretación bastante amplia del modo con el cual vienen tratadas las informaciones.

El producto creado está dotado de un nivel de transparencia tal que lo hace independiente del tipo de ambiente en el cual debe operar y de las metodologías de tratamiento de las informaciones.

Para facilitar la comprensión de las diferencias entre los varios métodos de tratamiento de los archivos informáticos, en la tabla a continuación vienen indicadas las terminologías en uso de tales ambientes.

TERMINOLOGIA COMPARATIVA

Relation DB	Data File System	Mathematics
Table	File	Relation
Column	Field Label	Attribute
Column Value	Field	Attribute Value
Row	Record	Tuple
Ext.Data Type	Data Type	Domain

Definición de los atributos de la Sub Entidad

Las reglas para la definición de las sub entidades que vienen indicadas en este banco integran los conocimientos del tipo informático que

Plataforma Informativa

Cuando se habla de plataforma informática se entiende el conjunto de computadoras y sistemas operativos que constituyen el sistema informativo. Las características de las diferentes plataformas exigen el respeto de standards predefinidos mediante los cuales está consentido de garantizar la operatividad a aplicaciones que han sido desarrolladas bajo los requisitos precisos.

La metodología de Ingeniería de Software aplicada y el respeto de precisos y rigurosos principios matemáticos, ha consentido desarrollar un sistema capaz de garantizar la portabilidad y la integridad a través de las fases de evolución de las diferentes plataformas informáticas.

Las versiones del producto actualmente, cubren las principales plataformas que son: Main Frames, Mini Computers e Personal Computers. En particular se trata de versiones donde los programas son prácticamente los mismos para las tres categorías. Donde, las interfaces necesarias para operar en el

ambiente de las plataformas específicas, vienen generadas en automático de programas contruidos para este fin. A continuación se indican las principales plataformas, a las cuales vienen referidos los lenguajes y sistemas operativos de posible uso.

PLATAFORMAS INFORMATICAS

MAINFRAME	MINI	PERSONAL
Cobol/2	Cobol/2	Cobol/2
"C" Language	"C" Language	"C" Language
PL/1	OS/400	PL/1
MVS/XA,ESA	Unix-Aix	OS/2 (FBSS)
Unix	VMS	Unix
VM	Oracle	MS/DOS
DB/2	SQL/400	Oracle-Ingress
IMS-DL/1	TCP/IP	Token Ring
CICS	X-Windows	Windows

Lenguajes

Los lenguajes de programación empleados cubren la mayor parte de los standards existentes actualmente en el mercado. Se trata de lenguajes que garantizan un alto nivel de optimización de los recursos informáticos disponibles.

En la tabla relativa a las plataformas informáticas están indicados los lenguajes utilizados en las versiones actuales. En la fase de generación del sistema vienen pedidas las informaciones relativas al tipo de lenguaje escogido y el ambiente informático donde se debe operar.

String

Por String se entiende una serie de letras, registradas en forma continua al interno de un mismo campo.

La delimitación del string inicia in presencia del primer caracter significativo, a partir de la primera posición de un campo y termina en presencia de una espacio o de un caracter especial. La operación se repite en todas las eventuales informaciones sucesivas y se termina al final del campo.

A continuación viene desarrollado un ejemplo en el cual vienen indicados los strings aislados.

ARCHIVO ORIGINAL

Razón Social	Dirección	Localidad
Mariano Rossani	Via Roma	Città di Cast.
Magazzini Bosco	Via Roma	C. di Castello
Rossani Mariano	Pza Bra	Battaglia T.
Bosco Mag.	Via Roma	Città di Castello
Mariano Rossani	Piazza Bra	Battaglia Terme
Rossani Mariano	Pza Bra	Battaglia T.
Magazzini Bosco	Borgo Real	Palo Vicentino
Mag. Bosco	Real Borgo	Palo Vic.
Rossani Mariano	Via Roma	Città Castello
Bosco Mag. Srl	V. Roma	Città di Castello'

STRINGS VALIDOS

Mariano	Rossani	Bosco	Roma	Bra
Borgo	Real	Città	Castello	Cast
Battaglia	Terme	Palo	Vicentino	

STRINGS ELIMINADOS

Magazzini	Mag.	Srl	Via	Piazza
Pza	v	C	di	Vic

Como indicado anteriormente, los strings toman son valorizados con una parte de un campo. Estos pueden presentarse invertidos y ser considerados solo una parte como validos. Los strings eliminados son considerados tales en cuanto forman parte de elementos comunes tratados más adelante en los puntos **Particelas Atributos, Puntuación, Tipología Empresarial etc.**

Atributos

Se llaman así los elementos de tipo generalizado los cuales vienen excluidos del análisis o bien, según las exigencias, tratados de acuerdo a reglas que pueden ser introducidas en el sistema en forma de guía.

En el archivo "Atributos" vienen suministrados una serie de elementos que pueden ser de utilidad en fase de instalación inicial. Los atributos nuevos pueden ser evidenciados a través del análisis de un congruente número de datos **anagráficos** y agregados, en forma guiada, en el sistema instalado. Tratándose de elementos generalizados, estos pueden ser modificados in fase de instalación o sino excluidos de la tabla en forma provisional o definitiva.

A continuación algunos ejemplos, para facilitar la comprensión del concepto de atributo definido en el sistema.

Abbigliamento	Acciaieria	Agenzia	Amministrazione
Associazione	Avvocato	Bar	Birreria
Carrozzeria	Cooperativa	Dist.	Distribuzione
Egregio	Ferramenta	Grafica	Impianti
Ingegnere	Laterizi	Latteria	Magazzino
Nastrificio	Pasticceria	Pizzeria	Ricambi
Salumificio	Sanitaria	SpA	Spett.
Termica	Tessuti	Utensili	Vetreria

Los atributos definidos para facilitar la terminología en este sistema pueden ser agrupados por tipologías. Se trata entonces de una posibilidad que el usuario puede aplicar según las reglas propias y que el sistema mantiene en modo automático.

La referencia a la tipología definida por parte del usuario viene a constituir, en el caso venga aplicada, una parte integrante de la clave atribuida por parte del sistema.

Control de los elementos anagráficos

En el presente trabajo monográfico el campo de acción de los productos de correlación y normalización son los datos **anagráficos** o también llamado Base Datos Clientes. En las tablas a continuación se puede notar una amplia gama de elementos con el tipo de control que estos pueden exigir para garantizar una congruencia requerida por parte de las aplicaciones bancarias.

De la evidenciación de los elementos contenidos en las tablas propuestas es fácil intuir la extensión del sistema tratado. En realidad, los elementos como por ejemplo el **CAB**, actividad económica, sexo y otros, pueden ser derivados de los archivos **anagráficos** o sino introducidos con los métodos tradicionales. Los dos métodos de igual forma son visto por el sistema.

Tipología Empresarial

Se trata de los elementos relativos a los diferentes tipos de Empresas en la gestión comercial italiana e internacional. La tabla que viene denominada **Atributos** tiene la capacidad de almacenar cualquier elemento que este dentro de esta casística. Los elementos introducidos solamente se refieren a la tipología italiana. La tipología internacional puede ser considerada en

cualquier momento e introducidas en el sistema por medio de las funciones específicas para tal fin.

En la siguiente tabla vienen indicados algunos elementos incluidos en la tabla Atributos.

TIPOLOGIA EMPRESARIAL

AS	CO	DI	EPG	IC
NR	PF	SAA	SCI	SCL
SDF	SM	SNC	SPA	SRL

Puntuación

Por puntuación se entiende la presencia de puntos, doble puntos, punto y coma y comas. En general, ha sido posible encontrar que se tratan de elementos que, al interno de un campo o de un string, no presentan. para efectos de relación, un nivel de significativa importancia.

Debido a verificaciones específicas y a la aplicación de reglas aptas a determinar la conveniencia de considerar la puntuación, ha sido igualmente posible notar la tendencia a eliminar los signos de puntuación al interno de las siglas o de nombres de Empresas. Ejemplo I.B.M. = IBM o sino F.I.A.T. = FIAT, etc.

Como consecuencia de tal tendencia, ha sido posible verificar que la presencia de la puntuación representa solamente una ocasión para limitar las posibilidades de agrupamiento de elementos símiles. Se trata por esto de una condición que, si no es aplicada con ciertas reglas, puede constituir un vínculo de congruencia y de uniformidad en la registración de los elementos **anagráficos**.

Caracteres especiales

Análogamente a la puntuación, los caracteres especiales por su naturaleza, constituyen una ocasión para limitar en parte la eficiencia del análisis y la uniformidad de explicación de las informaciones tratadas. Los caracteres utilizados en este sistema son indicados en la Tabla caracteres, y las reglas para su tratamiento son análogas a aquellas aplicadas en la parte relativa a la puntuación.

TABLA CARACTERES

!	“	#	\$	%	&	'	0
*	+	-	?	<	=	>	?
@	[\]	^	_	°	È
ª	º	¿	®	½	¼	«	»
¡	À	Á	Â	©	+	¢	¥
ã	ö	ð	é	ë	ß	ö	ö

Particelas

Se trata de **Particelas** que representan, en algunos casos, la parte integrante de nombres y apellidos italianos y demás. Se ha notado que, a efectos de análisis, no representan una condición limitante y de consecuencia no vienen considerados para la determinación de las claves de correlación.

A efectos de normalización, estas **particelas**, pueden ser descodificadas según reglas a definir o sino, por ejemplo, ser corregidas a caracteres mayúsculos o viceversa. En cuanto a la generación de la tabla relativa, esta

actividad viene realizada mediante reglas y modos previstos por la tabla "Atributos".

TABLA PARTICELA

di	del	delle	d'	dal	dei	per	in	f.lli
figli	c/o	gli	lo	la	le	il	tra	fu
al	alle	su	sul	sullo	degli'	degli	agli	gli

Caratteristiche

Con el termino **caratteristiche** se entiende aquellos elementos que caracterizan la guía de calles italianas y los nombres asociados al mismo. Estos presentan las mismas particularidades de la puntuación, de las **particelas** y de los caracteres especiales. Dado el tipo de similitud el tratamiento resulta análogo a los mismos. A continuación viene indicada la tabla utilizada en fase de instalación inicial y una parte de los datos contenidos en la misma.

Como **caratteristiche** vienen tratadas también los elementos que conforman la guía de calles de la ciudad.

TABLA CARACTERISTICAS

San	Santo	Santo	SS	dx	sx	rosso	nero
Via	Strada	Piazza	Borgo	Zona	Largo	Viale	Calle

Tablas usadas en el Sistema

1. Nombres propios de personas
2. Descripciones en general
3. Preposiciones y artículos
4. Descripciones de lugares
5. Atributos de números cívicos
6. Tipología de las localidades
7. Tipología de los meses
8. Tipología de los números
9. Títulos académicos
10. Elementos religiosos
11. Tipologías calles

TABLA 1. Nombres propios de personas

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	ABELARDO		
2	ABIR		
3	ABRAMO		
4	ACHILLE		
5	ADAD		
6	ADALBERTO		
7	ADALGISA		
8	ADAMO		
9	ADDOLORATA		
10	ADELAIDE		
11	ADELE		
12	ADELINA		
13	ADELMO		
14	ADEODATO		
15	ADOLFO		
16	ADRIANO		
17	AGAMENNONE		
18	AGATA		

TABLA 2. Descripciones en general

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	ABBIGLIAMENTI	ABBIGLIAMENTI	
2	ABBIGLIAMENTO	ABBIGLIAMENTO	
3	ACCESSORI	ACCESSORI	
4	ACCIAIERIE	ACCIAIERIE	
5	AG	AGENZIA	
6	AGENCY	AGENZIA	
7	AGENZIA	AGENZIA	
8	AGRICOLA	AGRICOLA	
9	AGRICOLE	AGRICOLE	
10	ALBERGO	ALBERGO	
11	AMM	AMMINISTRAZIONE	
12	AMMINISTRAZIONE	AMMINISTRAZIONE	
13	ASSOC	ASSOCIAZIONE	
14	ASSOCIAZIONE	ASSOCIAZIONE	
15	AZ	AZIENDA	
16	AZI	AZIENDA	
17	AZIENDA	AZIENDA	
18	AZIENDE	AZIENDA	
19	CASELLA	CASELLA POSTALE	
20	CASELLAPOST	CASELLA POSTALE	

TABLA 3. Preposiciones y artículos

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	AD	AD	
2	AI	AI	
3	AL	AL	
4	ALL	AL	
5	ALLA	AL	
6	ALLE	AL	
7	ALLI	AL	
8	ALLO	AL	
9	CA	CA	
10	CON	CON	
11	DA	DA	
12	DAL	DAL	
13	DALLA	DAL	
14	DALLE	DAL	
15	DALLI	DAL	
16	DALLO	DAL	
17	DE	DE	
18	DEG	DEG	
19	DEGL	DEG	
20	DEGLI	DEG	

TABLA 4. Descripciones de lugares

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	ACQ	ACQ	
2	ACQUEDOTTO	ACQ	
3	AEROP	AEROP	
4	AEROPORTO	AEROP	
5	AEROT	AEROT	
6	AEROTRASPOSTI	AEROT	
7	AMM	AMM	
8	AMMIRAGLIO	AMM	
9	ARCD	ARCID	
10	ARCID	ARCID	
11	ARCIDUCA	ARCID	
12	ARCIV	ARCIV	
13	ARCIVESCOVO	ARCIV	
14	ARM	ARM	
15	ARMATA	ARM	
16	ART	ART	
17	ARTIG	ARTIG	
18	ARTIGIANA	ARTIG	
19	AUS	AUS	
20	AUSILIATRICE	AUS	

TABLA 5. Atributos de números cívicos

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	APP	APP	
2	APPARTAMENTO	APP	
3	ASC	ASC	
4	ASCENSORE	ASC	
5	FABB	FB	
6	FB	FB	
7	INT	INT	
8	INTERNO	INT	
9	IS	IS	
10	KM	KM	
11	LOT	LT	
12	LOTTO	LT	
13	LT	LT	
14	NO	NRO	
15	NR	NRO	
16	NRO	NRO	
17	NUM	NRO	
18	NUMERO	NRO	

TABLA 6. Tipología de las localidades

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	FARZ	LOC	
2	FNE	LOC	
3	FR	LOC	
4	FRANE	LOC	
5	FRAZ	LOC	
6	FRAZIONE	LOC	
7	FRAZNE	LOC	
8	FRNE	LOC	
9	FRZ	LOC	
10	FRZNE	LOC	
11	LOC	LOC	
12	LOCAL	LOC	
13	LOCALITA	LOC	
14	LOCTA	LOC	
15	LTA	LOC	

TABLA 7. Tipología de los meses

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	AGO	AGO	
2	AGOSTO	AGO	
3	APR	APR	
4	APRILE	APR	
5	DIC	DIC	
6	DICEMBRE	DIC	
7	FEB	FEB	
8	FEBBRAIO	FEB	
9	GEN	GEN	
10	GENNAIO	GEN	
11	GIU	GIU	
12	GIUGNO	GIU	
13	LUG	LUG	
14	LUGLIO	LUG	
15	MAG	MAG	
16	MAGGIO	MAG	

TABLA 8. Tipología de los números

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	1	1	
2	10	10	
3	11	11	
4	12	12	
5	13	13	
6	14	14	
7	CINQUE	5	
8	DECIMO	10	
9	DICIANNOVE	19	
10	DICIANNOVESIMO	19	
11	DUE	2	
12	II	2	

TABLA 9. Títulos académicos

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	AMM	AMM	
2	AMMINISTRATORE	AMM	
3	ARCH	ARCH	
4	ARCHITETTO	ARCH	
5	AVV	AVV	
6	AVVOCATO	AVV	
7	COMM	COMM	
8	COMMENDATORE	COMM	
9	DOTT	DOTT	
10	DOTTORE	DOTT	
11	DOTTORESSA	DOTT	
12	DR	DOTT	

TABLA 10. Elementos religiosos

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	S	SAN	
2	SAINT	SAN	
3	SAN	SAN	
4	SANT	SAN	
5	SANTA	SAN	
6	SANTI	SAN	
7	SANTISSIMI	SAN	
8	SANTISSIMO	SAN	
9	SANTO	SAN	

TABLA 11. Tipologías calles

ITEM	STANDARD	USUARIO	OBSERVACIONES
1	ACC	ACC	
2	ACCESSO	ACC	
3	ACCO	ACCO	
4	ACCORCIATO	ACTO	
5	ACCORCIATOIA	ACTO	
6	ACTO	ACTO	
7	AER	AER	
8	AEROP	AER	
9	AEROPORTO	AER	
10	AEROSCALO	AERS	
11	AERS	AERS	
12	ASC	ASC	
13	ASCENSORE	ASC	
14	CLIV	CLIV	
15	CLIVIO	CLIV	
16	CLIVO	CLIV	
17	ISO	ISO	
18	ISOLA	ISO	
19	ISOLATO	ISO	

BIBLIOGRAFIA

- 1. Concetti in Ingegneria del Software, Mondadori Informatica - Italia 1996**
- 2. Teoria di Normalizzazione, Mondadori Informatica - Italia 1992**
- 3. Algoritmi di Ricerche alfanumeriche, Edizioni Finson - Italia 1995**
- 4. Harling John. " Simulation Techniques in Operations Research - A Review
", Operations Research.**
- 5. James Martin. " Base de Datos Relacionales ". Edición 1990.**