# **Feature Requirements Document**

Feature Name:Time independent ClusteringRequest Date:8/25/99Version:3Completed by:Leon Guzenda

## **Description of the Problem:**

In some systems data is collected over time but needs to be clustered together. The current clustering mechanism is such that data is stored sequentially in pages. This means that over time an object and its related objects can be scattered over many pages. For example an application could create a person object related to several other objects (e.g. insurance policies, X-rays, items). Over time new objects will be created that are related to the original person object, but these will be created physically separate (different pages) to the original object. This can lead to performance degradation, for instance instead of all objects clustered together on one or two pages the objects are spread over many pages. Even during database load all data that should be clustered together will not be available, e.g. loading data from any different systems.

## **Description of the requested feature:**

The ability to cluster object groups on the minimum number of pages determined only by physical space requirements, independent of time of creation. This requires that the first object is stored on a new page and all subsequent objects belonging to the same object group are clustered together with the first object. If the page runs out of space, an overflow page will store the new objects. Only objects belonging to the same object group will be stored on the first page and its overflow pages.

## Part of an optional feature or does it require another feature? If so, which one?

None

## How is this problem being solved now, and why isn't that acceptable?

Currently application code has to be written to pre-allocate space in pages, leaving storage management to the application programmer.

## Which languages must support this capability?

All.

## Which platforms must be supported?

All.

## Do any competitors already have this capability?

Not that I know of

## Customers who require this capability:

Cuna, AWD, TRW, any other enterprise (e.g. insurance, health care, customer care, financial, where the data involves people interactions)

## Revenue at risk or which could be won:

## When is this required?

July '99 for testing before the Cuna live database load.

# **Revision History:** Version 1 7/6/98 Brian Clark Version 2 1/5/99 Leon Guzenda Version 3 8/25/99 Leon Guzenda