M E M O R A N D U M

To: David Marwell

From: Jeremy Gunn

cc: Sheryl Walter, Phil Golrick, Kevin Tiernan

Date: February 27, 1995

Re: HSCA database

_____

The HSCA "postponed" database appears to be in disarray.

1.   Misidentification of Agency file numbers.

HSCA documents physically are arranged by Agency file number.   Thus the easiest way for Analysts to review the HSCA documents will be by HSCA file number.   Unfortunately, there appears to have been little or no quality control over the entry of these file numbers.

I asked Chet to create a Lotus Notes "view" for Agency file number (see View: Agency number, RIF, Status etc).   As you will be able to see, some documents were entered with six digit codes and some were entered with seven digit codes.   Thus documents that physically are in sequential order are not in sequential order in our database.   Some of the documents are given altogether different numbers (*e.g.,* FBI serial numbers).

These mistakes should be corrected.

2.   Referrals.

According to our master postponed database, there are a total of 4,462 HSCA records.   (Because of database entry problems, there are obvious additional records that should be added to the HSCA number.   For example, there are 20 records misidentified as "HCSA" and 2 misidentified as "BSCA.")

The HSCA database identifies 4,130 of these records as having been referred to other agencies.   According to Kris Wilhelm and Matt Fulgham, however, the most accurate listing of referred documents is actually the NARA looseleaf binders that contain handwritten photocopies of the HSCA RIFs.   NARA uses these photocopies as their principal tracking device to determine which records were sent to which agencies and whether the agencies returned the records after review.   In theory, therefore, all records that NARA referred should have a RIF hardcopy in the binders.

In order to make sure that our database is as accurate as possible, we obtained photocopies of the NARA RIFs from the binders.   Last week the Analysts entered the data from the photocopies into the HSCA database and recorded to which agency or agencies the records were sent.

I asked Chet to create a Lotus Notes "view" for the HSCA database that would allow me to determine the number of records that were referred to each agency.   (See View: Referred to; RIF.)

Unfortunately, **3,251** of the HSCA postponed documents do not have any hardcopy record of having been referred to any agency (see category 31 "not categorized").   Thus we do not know for certain whether these records have in fact been sent to the agencies or, for that matter, to which agency

-- presuming that they were referred.   This is a particular problem where the entries for "Agency" and "Originator" are inaccurate -- as is frequently the case.   (*E.g.,* sometimes the "originator" is identified as the person whose deposition was taken or the person who authored a document. Additionally, the HSCA is often identified as the "Agency" when there is a deposition of a CIA or FBI employee.)

We have some reason to believe that many of the "referred" documents may in fact be open.   We will be doing some spot checking to see if we can get a handle on the approximate numbers.

3.   Proposals.

I presume that the problems identified above will be typical of what we will face during the next three years.   Accordingly, I would like to make the following proposals:

First, we should obtain the latest *postponed* database updates that Steve Tilley has in his possession.   Apparently, Steve has a significant amount of information that he has not been able to add to his system.

Second, we should get the updated information into our postponed database ASAP.

Third, after we get the updated information, we should cut ourselves off from Steve's postponed database and make all future corrections ourselves.

We need to have the best and most current information available for postponed documents.

Fourth, once we have the updated information, we should start making clean-up corrections to the database.   The HSCA Agency file numbers is a good place to start.

Fifth, we should consider dividing responsiblity with Steve for maintaining the databases.   It makes sense that we track, correct, and update all postponed documents and that Steve be responsible for correcting RIFs for all open documents.   We can periodically exchange databases from him. Regardless of whether we formalize this suggestion, however, I do believe that we should take the initiative to develop the most accurate postponed database that we can.