

**A Web-Based Database
of
CIA Declassified Documents on the Vietnam War**

Vinh-The Lam and Darryl Friesen

(Published in ***Online***, v. 28, no. 4, 2004, p. 31-35)

Abstract:

US Government classified documents generated during the Vietnam War years are now being declassified and made available for educational and research purposes. This database includes only those documents produced by the US Central Intelligence Agency. It provides an in-depth indexing of the documents' contents and, when possible, have links to full-text documents available online at the Website of the Virtual Vietnam Archive. The database is designed to offer both simple and advanced search capabilities.

Keywords: US Government Declassified Documents -- US Central Intelligence Agency -- CIA -- Online Database -- Web Database

Introduction

During the Vietnam War years (1960-1975), the United States Government has generated a large volume of classified documents. The declassification of these documents started with Executive Order No. 11652 signed by President Richard Nixon in 1972.¹ Thousands of these documents, formerly classified as **Confidential**, **Secret**, and **Top Secret**, are being declassified, made public, and available for educational and research purposes. They are put on microfiche and published by Primary Source Microfilm as *Declassified Documents Reference System* (DDRS). The microfiche are abstracted, indexed and published in a bi-monthly periodical titled *Declassified Documents Catalog* (DDC). The DDC is now also published as a CD-ROM by the Gale Group and the DDRS is also available through subscription on the INTERNET at this Website: <http://www.ddrs.psmedia.com/> Recently, the Vietnam Center of Texas Tech University in Lubbock, Texas, through its Virtual Vietnam Archive (VVA), available at: <http://www.vietnam.ttu.edu/virtualarchive/>, also provides access to a large number of fulltext declassified documents.

This database is the result of a Sabbatical Leave research project approved and supported by the University of Saskatchewan, Canada. It includes only declassified documents created by the US Central Intelligence Agency (CIA). It provides an in-depth indexing of the CIA declassified documents and, where possible, also provides a link to the full-text documents available at the VVA.

Database Structure

Each record in the database contains the following fields:

Record Number: Automatically created by system

Title: Title of document

Date of Creation: Date document was created

Date of Declassification: Date document was declassified

Type of Document: Type of document, e.g., Report, Memorandum, Cable, etc.

Level of Classification: Level of classification of document before it was declassified; only four terms will be used: CONFIDENTIAL, SECRET, TOP SECRET, and NOT GIVEN

Status of Copy: Status of copy of document ; only two terms will be used: ORIGINAL, and SANITIZED

Pagination: Number of pages, and illustrations, e.g. maps

Abstract: Abstract of contents of document; taken mostly from the CD-ROM published by the Gale Group

Indexing Terms: Controlled vocabulary (words, phrases) describing topics presented in document

DDRS Location: Document identifier showing location of document in the Declassified Documents Reference System

Link to Full-text: If available, URL of document available full-text at the Website of the VVA

Document Indexing and Database Contents

The main reason for this database to be created is the lack of in-depth indexing that makes searching the DDC unsatisfactory. The very detailed indexing provided by the Carrollton Press for the *Declassified Documents Retrospective Collection*, published in 1976 was abandoned when Carrollton Press began publishing the *Declassified Documents Quarterly Catalog*, which preceded the DDC. Research Publications adopted this practice for the DDC. When Primary Source Microfilm replaced Research Publications as publisher of DDC, it continued this practice. As a result, a very limited number of indexing terms are used in the DDC as follows:

Vietnam
Armed Forces
Foreign relations with ---
Politics and government
Religion

- Vietnam, North
 - Commerce
 - Foreign relations with ---
 - Military policy

- Vietnam, South
 - Armed forces
 - Commerce
 - Commerce with ---
 - Economic conditions
 - Foreign relations with ---
 - Politics and government
 - Religion
 - Social conditions

- Vietnamese Conflict, 1961-1975
 - Campaigns
 - Missing in action
 - Peace negotiations
 - Prisoners of war

The topical searches such as searches for personal names, place-names, names of operations/battles, titles of US and/or Vietnamese government projects/programs, etc., which would be very useful for Vietnam War scholars/researchers, are impossible.

The authors decided, therefore, to provide an in-depth content analysis of the documents. Full-text documents were analyzed thoroughly page-by-page so that names of people (e.g., politicians, military leaders, etc.), operations/battles, military units (US, Allied, South Vietnamese, North Vietnamese, Viet Cong divisions, regiments, battalions, etc.) projects/programs, place-names (provinces, cities, towns, valleys, mountains, rivers, etc.) could be picked up and used as indexing terms.

For example, a search for the most important Communist offensive of the war, which was called the **Tet Offensive**, would retrieve the following:

Declassified CIA Documents on Vietnam War - Microsoft Internet Explorer

Address: <http://library.usask.ca/vietnam/index.php>

University of Saskatchewan Library

UofS Library > Declassified CIA Documents on the Vietnam War [select your branch]

Declassified CIA Documents on the Vietnam War

We hereby acknowledge the permissions granted by the Gale Group to include in this Database the abstracts provided in the Declassified Documents Reference System CD-ROM and by the Texas Tech University's Virtual Vietnam Archive to provide the links for fulltext documents.

Your search for **tet offensive** matched **63** documents
 Documents 1 to 20 are shown below
 View documents: [1 to 20] Search: Go [Advanced search...](#)

SQL: SELECT * FROM DeclassifiedDocuments WHERE CONTAINS(*, "tet" AND "offensive")

ID	Document
01061	Second phase of Viet Cong Tet Offensive (May 9, 1968) Intelligence information cable; CONFIDENTIAL; SANITIZED Declassified: October 10, 1997
01091	Situation report as of 24 hours local time, 11 February 1968 (February 11, 1968) Cable; SECRET; SANITIZED

Start | Internet | 2:46 PM

A search for the famous US 101st Airborne Division would retrieve the following:

Declassified CIA Documents on Vietnam War - Microsoft Internet Explorer

Address: <http://library.usask.ca/vietnam/index.php>

University of Saskatchewan Library

UofS Library > Declassified CIA Documents on the Vietnam War [select your branch]

Declassified CIA Documents on the Vietnam War

We hereby acknowledge the permissions granted by the Gale Group to include in this Database the abstracts provided in the Declassified Documents Reference System CD-ROM and by the Texas Tech University's Virtual Vietnam Archive to provide the links for fulltext documents.

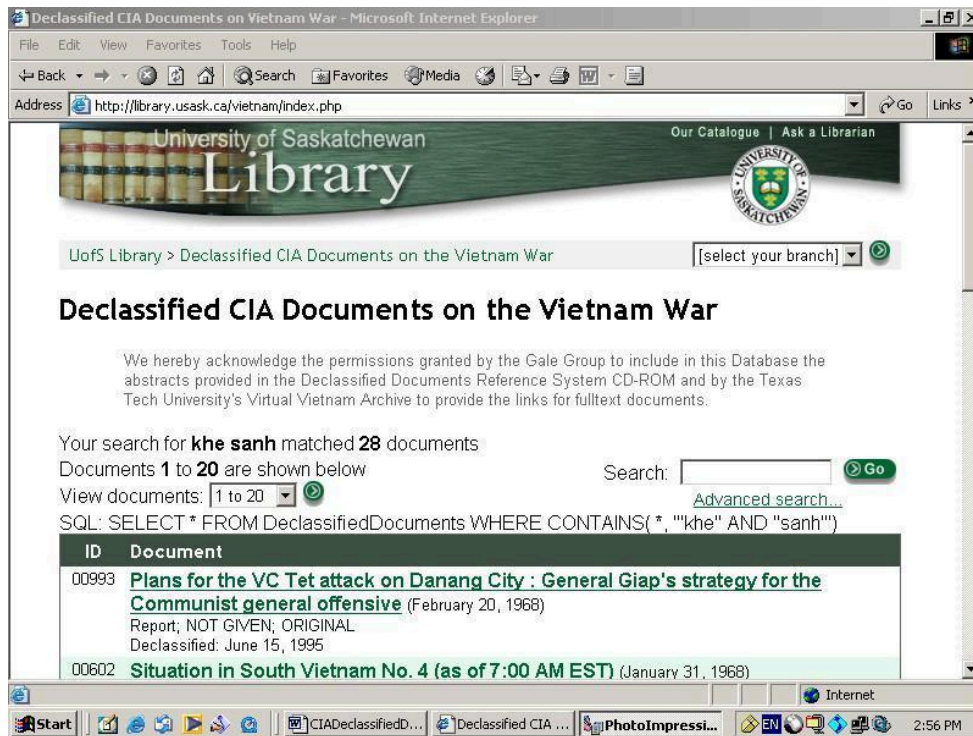
Your search for **us 101st airborne division** matched **6** documents
 Search: Go [Advanced search...](#)

SQL: SELECT * FROM DeclassifiedDocuments WHERE CONTAINS(*, "us" AND "101st" AND "airborne" AND "division")

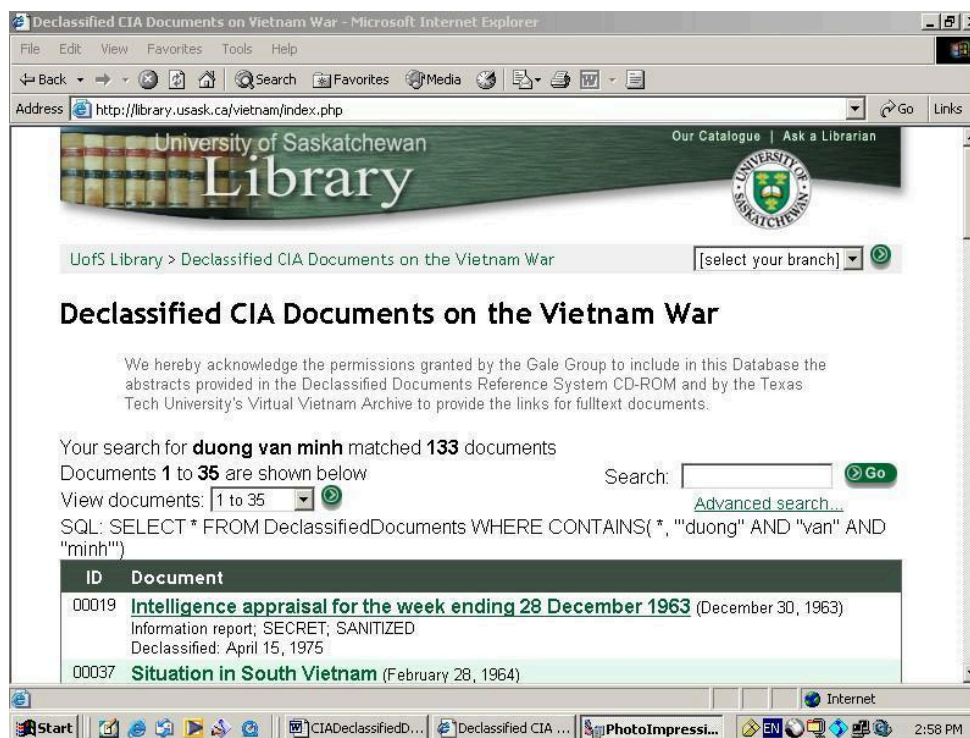
ID	Document
00475	Situation in South Vietnam (March 2, 1966) Weekly report; SECRET; SANITIZED Declassified: August 20, 1980
00941	Situation in South Vietnam (April 20, 1966)

Done | Internet | 2:53 PM

A search for Khe Sanh, location of the bloodiest battle between the US Marines and the North Vietnamese divisions, would retrieve the following:



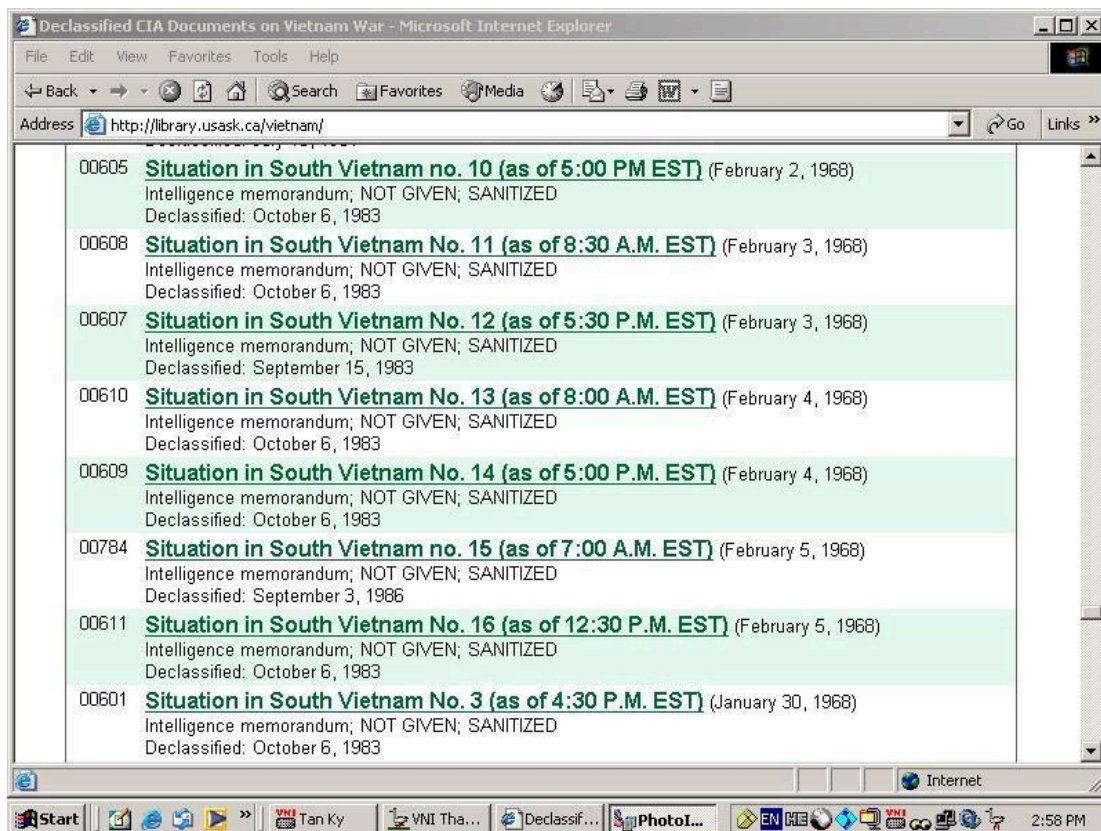
A search for General Duong Van “Big” Minh, leader of the military coup that overthrew the Ngo Dinh Diem government on November 1st, 1963, would retrieve the following:



In addition to the in-depth indexing, the authors also tried to achieve consistency for indexing terms assigned to records throughout the whole database in order to maximize retrieval. It was decided to provide personal names in the non-inverted form, e.g., Duong Van Minh instead of Minh, Duong Van, or Robert McNamara, not McNamara, Robert. Since one of the co-authors is of Vietnamese origin, wrongly spelled Vietnamese names in documents were detected and corrected. South Vietnamese government program titles were translated into English. Sometimes both English and Vietnamese forms of the program titles, if already familiar within the Vietnam War research community, were used as equivalent indexing terms, e.g. ***Returnee Program*** and ***Chieu Hoi Program***. When the database was populated with about 500 records, a thorough review and revision of all indexing terms was carried out to detect and correct typos, inconsistencies. A second review/revision was done when the database reached the 1,000-record level. The index now contains 3461 terms and its complete listing is 101-page long.

The database currently contains 1,080 records, 34% of which provide a link to the full-text documents available online at the VVA. The documents analyzed could go

from 1 page to a few hundred pages. They could be a Memo, a Telegram, a Report (weekly, monthly, etc.), a Situation Report (or SitRep), a Biographical Sketch, a National Intelligence Estimate (or NIE), a Special National Intelligence Estimate (or SNIE), a Research Study Report, etc. Sometimes, while an important event was occurring, e.g., the **Tet Offensive**, the CIA produced Intelligence memos on a daily or even hourly basis, as we can see in the following:



After the Johnson Administration decided to send combat troops to South Vietnam in 1965, the CIA produced weekly and monthly reports called ***The Situation in South Vietnam***, in which details on political, military, and economic situation of South Vietnam were given. In the Political Situation sections, the reports give detailed information of activities of the South Vietnamese Government, such as cabinet reshuffles, inauguration/development/changes of government programs/projects, and deliberations within the National Assembly. Also included are information on activities of political parties and their leaders, on rumors of possible coups, on

local/provincial/national elections, etc. In the Military Situation sections, the reports give detailed account of operations/battles engaging US, Allied, South Vietnamese, North Vietnamese, and Viet Cong units, as well as their casualties and weapon losses. The Economic Situation sections report important economic indicators, e.g., retail prices index (especially prices of rice and pork), weekly and monthly prices of gold and currency in the Saigon free market, etc. About 150 such reports are now included in the database and they are extremely useful for researchers who want to draw a chronological picture of South Vietnam during these war years, especially between 1965 and 1968. Another series of reports present monthly evaluation of the cost-effectiveness of **Operation ROLLING THUNDER**, which carried out the US sustained bombing of North Vietnam. Still another one details the level of North Vietnamese Army infiltration into South Vietnam. A close look at those reports, together with the NIEs and SNIEs on Vietnam will help database users understand how US policy on Vietnam was conceived and implemented. A large number of these declassified documents are sanitized, i.e., with source of information, and names of informants removed for protection purposes.

Database Design

Microsoft SQL Server 2000 was used as the database server. In addition to being an outstanding relational database server, the rich full-text search capabilities it offers made it an excellent choice for this project.

The database consists of a single table, although some normalization could have been done, especially with respect to the indexing terms. However, considering the few data elements in the table, the relatively small number of documents indexed, and the strength of the SQL server's search capabilities, a simple design was favored.

All of the columns in the table, with the exception of DocumentID, are variable length character data (`varchar`). DocumentID is an auto-incrementing integer value, managed by the SQL Server, and used as the primary key for the table. The CreationDate and DeclassificationDate fields were initially standard SQL `datetime`

data types, but had to be changed to character fields because of a bug in one of the underlying software components.

The following SQL statement was used to create the table in the database.

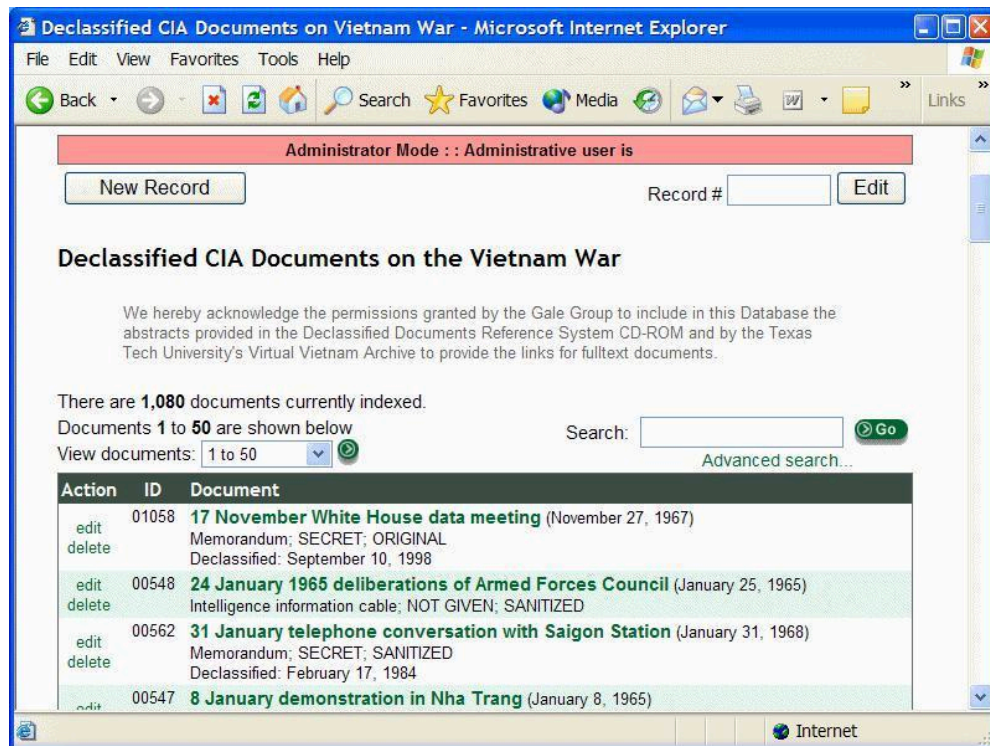
```
CREATE TABLE DeclassifiedDocuments ( DocumentID
    int IDENTITY (1, 1) NOT NULL , Title
    varchar(512) NOT NULL , CreationDate
    varchar(15) NULL , DeclassificationDate
    varchar(15) NULL , DocumentType
    varchar(254) NULL , ClassificationLevel
    varchar(50) NULL , CopyStatus varchar(10)
    NULL , Pagination varchar(100) NULL ,
    Abstract varchar(8000) NULL , Descriptors
    varchar(8000) NULL , DDRS_Location
    varchar(50) NULL ,
    URL varchar(254) NULL
)
```

All access to the database, including data entry and other administrative functions, is done using a web browser. The web-based user interface was written in the PHP programming language. PHP has experienced a rapid growth in popularity in recent years, due in part to its excellent handling of textual data (such as the data sent and received via web browsers and server) and database support.

The web server is a Sun UltraEnterprise 2 server running the Solaris 8 operating system, and a recent version of the Apache web server software. An open source product called FreeTDS allows the Unix web server to communicate with the Microsoft SQL Server directly using the TDS (Tabular Data Stream) protocol. TDS is the native protocol used by Microsoft and Sybase for their database products. Although still somewhat a fledgling product, FreeTDS is a workable solution for establishing connectivity between Unix machines and Microsoft or Sybase database servers.

Administrative functions – adding, modifying and deleting records – are also performed using the web browser. The administrative features are located in a secure area on the web server, and are password protected. The administrative interface closely resembles the public view, with the addition of links in both the brief and full records display which allow the document to be easily modified or deleted. A new record can be added by simply clicking the **New Record** button located at the top of the screen. In addition to being able to browse or search for records requiring modification,

a quick edit feature, located in the upper right, is available for documents whose Document ID number is known.



A simple web-based form is used for data entry and record editing.

Declassified CIA Documents on Vietnam War - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Mail Links

Document ID: 01058

Title: 17 November White House data meeting

Date of Creation: 11/27/1967 mm/dd/yyyy

Date of Declassification: 09/10/1998 mm/dd/yyyy

Type of Document: Memorandum

Level of Classification: SECRET

Status of Copy: ORIGINAL

Pagination, Illustration: 3 p.

Abstract: Memorandum for the record by George A. Carver, Jr., regarding the 11/17/67 White House meeting to coordinate data problems between Wash. (D.C.) and the U.S. Mission

Indexing Terms: ROBERT KOMER
WILLIAM JORDEN
RICHARD MOOSE

Declassified Documents Reference System Location: 1999-0689

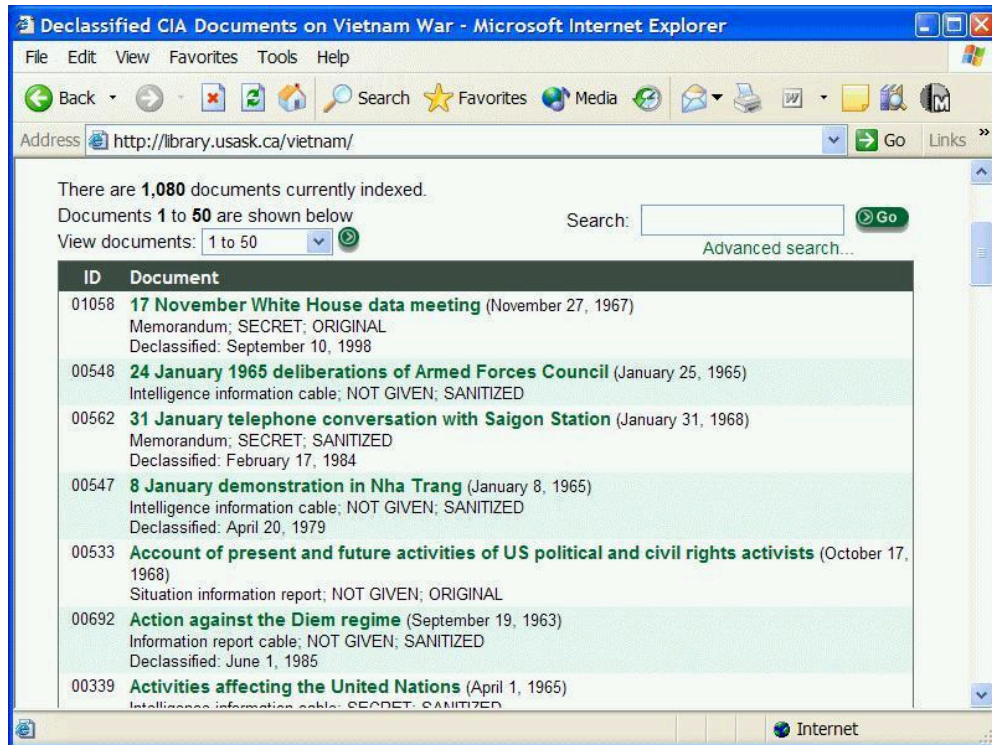
Link to Full Text:

Update Record

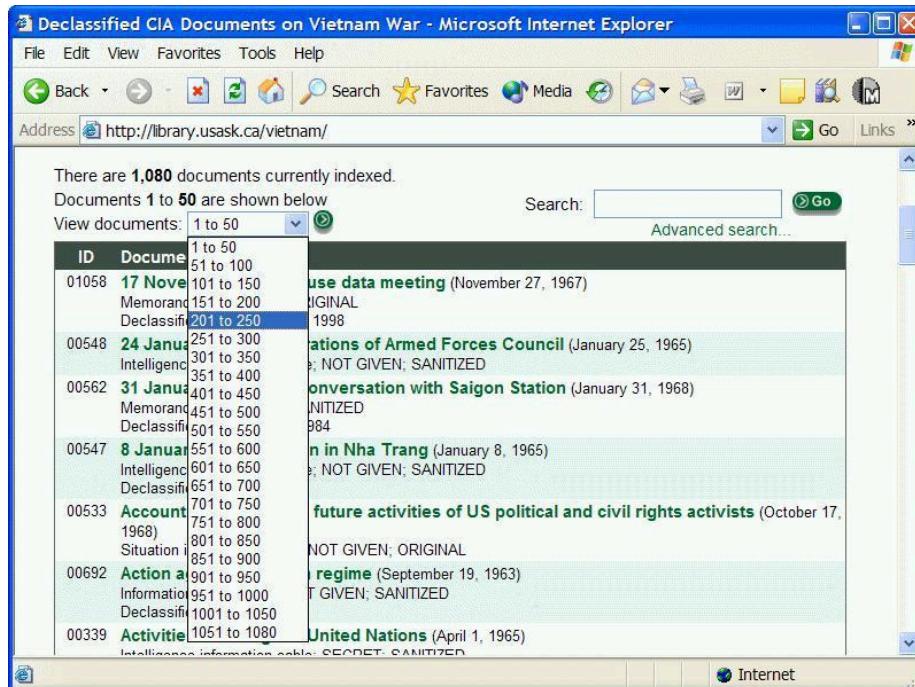
Internet

Database Navigation

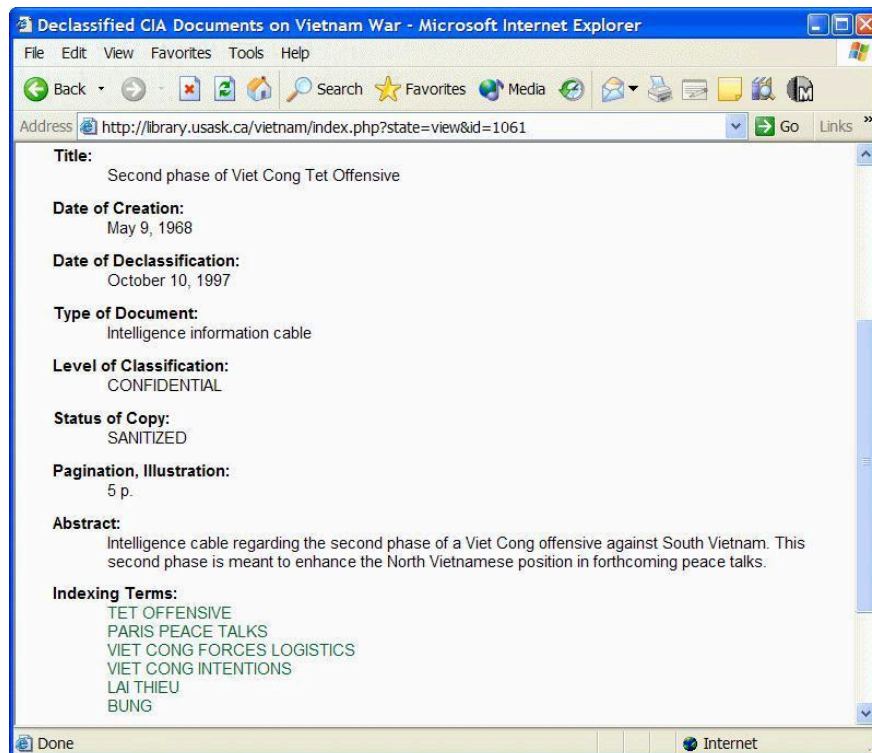
The default view is an alphabetical listing of all indexed documents, shown in a brief record format. Included in the brief citation is document title, creation date, declassification date, type of document, level of classification and status of copy.



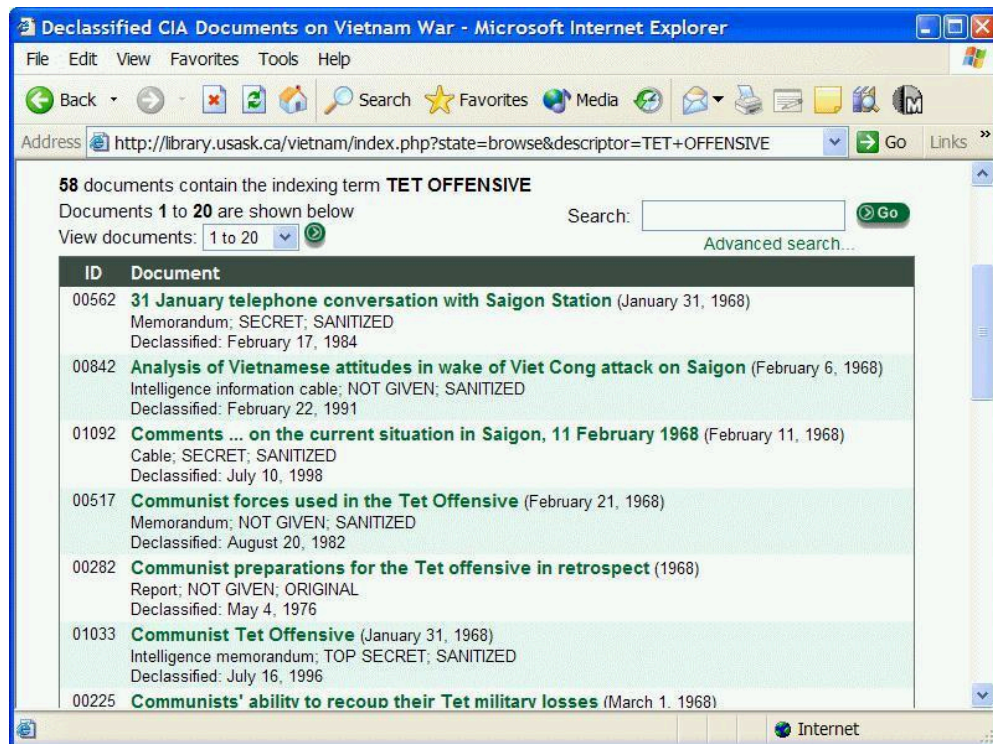
The number of documents displayed in the brief format is limited to 50 per page. A drop down menu provides easy access to all indexed documents, in groups of 50. Limiting the display in this manner, rather than listing all 1,080 documents at once, significantly decreases the time it takes a web browser to load the page, and from a usability point of view, increases the functionality of the database.



Clicking the title brings up a full record display for the selected document. Included in this view are fields not shown in the brief display, including pagination, abstract, indexing terms, DDRS location and URL.



The URL, if present, will link to the full text of the document in Virtual Vietnam Archive. Indexing terms are also hyperlinked, and clicking one will return all documents which share that indexing term.



Retrieval Mechanism

Microsoft SQL Server 2000 allows full-text indexes to be defined on selected columns in a table. This permits complex searches to be executed against any of the columns in the index, or all the columns at once. Boolean operators, phrase searching, word stemming, weighting, proximity searching and wildcard operators are all supported. Unlike traditional indexes defined on columns, SQL Server full-text indexes reside outside the database, on the server's local file system, and additional steps must be taken to populate them. Index population can be scheduled to occur at any time. In the case of this database, a full population and rebuilding of the index occurs once a week during off hours (5 a.m. Saturday morning), and an incremental population happens hourly during the period when data entry might normally occur (weekdays between 7 a.m. and 7 p.m.). This schedule ensures that the full-text index is up-to-date

with any additions or changes to database records. The document title, document type, level of classification, status of copy, creation date, declassification date, abstract and indexing term fields are all included in the full-text index.

The full-text index is utilized by both the simple keyword and an advanced searches. In the case of the simple keyword search, located conveniently at the top of almost every page, all terms entered by the user are joined with the Boolean AND operator and a search is performed across all fields in the full-text index.

The advanced search is quite powerful, allowing for more control both in terms of what is searched for and how, as well as the limits that are applied. Complex queries in which specific phrases are combined with a list of terms, all of which must appear, limited by classification level, copy status, full-text availability, and date range limits for both the document creation and declassification dates can be constructed. The following outlines a search for documents containing the exact phrase “**tet offensive**” created between January 1st and December 31st 1968, declassified before 1990 and in their original form.

Declassified CIA Documents on the Vietnam War

We hereby acknowledge the permissions granted by the Gale Group to include in this Database the abstracts provided in the Declassified Documents Reference System CD-ROM and by the Texas Tech University's Virtual Vietnam Archive to provide the links for fulltext documents.

Find results

with **all** of the words

with the **exact phrase**

with **at least one** of the words

Search

Classification Return documents whose classification level is

Status Return documents whose copy status is

Dates Return documents **created**

after Most common date formats are acceptable, and either of the start or end dates may be omitted. Dates are inclusive.

and before

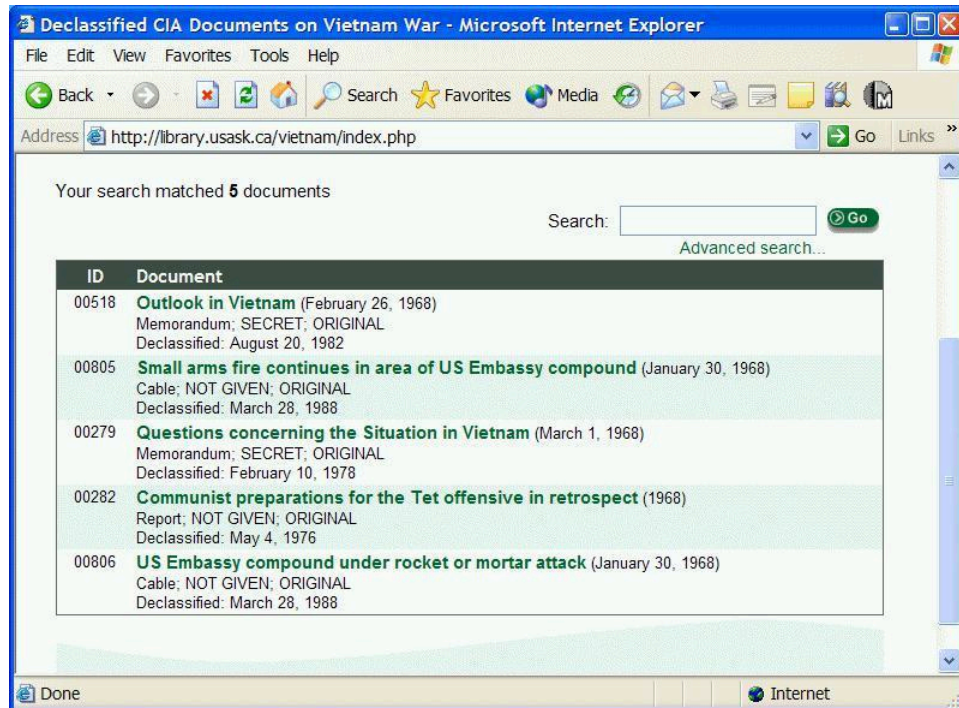
Return documents **declassified**

after

and before

Full text ☐ Only return documents that are available online

Refining the search using these advanced options decreases the result set from 46 (for a simple keyword search on **TET OFFENSIVE**) to 5.



The search terms themselves can be considered optional, and queries making use of just the limiting features acceptable. The following represents a legitimate search for all top secret documents, in their original form, whose full-text is available online.

Declassified CIA Documents on the Vietnam War

We hereby acknowledge the permissions granted by the Gale Group to include in this Database the abstracts provided in the Declassified Documents Reference System CD-ROM and by the Texas Tech University's Virtual Vietnam Archive to provide the links for fulltext documents.

Find results

with **all** of the words

with the **exact phrase**

with **at least one** of the words

Classification Return documents whose classification level is

Status Return documents whose copy status is

Dates Return documents **created**

after *Most common date formats are acceptable, and either of the start or end dates may be omitted. Dates are inclusive.*

and before

Return documents **declassified**

after

and before

Full text ☒ Only return documents that are available online

Conclusion

This online database was designed to provide an efficient tool for Vietnam War scholars/researchers to search for Declassified CIA Documents on various specific topics, with some possibility to retrieve full-text documents. The web-based user interface, written in the PHP programming language, provides users with an easy and smooth database searching, retrieval, and navigation. As CIA classified documents continue to be declassified, and with a firm commitment from the University of Saskatchewan Library, this database will continue to grow.

Notes

1.____ Morehead, Joe and Mary Fetzer. *Introduction to United States Government Information Sources*. 4th ed. Englewood, Colo. : Libraries Unlimited, 1992. p. 376.

Part of the Executive Order No. 11652 is accessible online at the following URL:

http://www.fas.org/sgp/eprint/legacy_appendix.html

