

**IN THE UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF TEXAS
TEXARKANA DIVISION**

THE STATE OF TEXAS,	§	CIVIL NO.: 5:96-CV-0091
Plaintiff,	§	
	§	
VS.	§	
	§	JUDGE: DAVID FOLSOM
THE AMERICAN TOBACCO	§	
COMPANY; R. J. REYNOLDS	§	
TOBACCO COMPANY;	§	MAGISTRATE JUDGE:
BROWN & WILLIAMSON TOBACCO	§	WENDELL C. RADFORD
CORPORATION; B. A. T.	§	
INDUSTRIES, P. L. C.; PHILIP	§	
MORRIS, INC.;	§	
LIGGETT GROUP, INC.; LORILLARD	§	
TOBACCO COMPANY, INC.; UNITED	§	
STATES TOBACCO COMPANY; HILL	§	
& KNOWLTON, INC.; THE	§	
COUNCIL FOR TOBACCO.	§	
RESEARCH-USA, INC. (Successor to	§	
Tobacco Institute Research Committee);	§	
and THE TOBACCO INSTITUTE, INC.	§	
Defendants.	§	

**PLAINTIFF'S MOTION TO COMPEL
DOCUMENT DATABASES IN A USABLE FORM**

Plaintiff, the State of Texas, respectfully requests this Court compel Defendants to produce certain computer diskettes containing document indices and other databases in a useable form.

1. BACKGROUND

On November 13, 1996 this Court ordered all parties to produce computerized indices for all documents produced to other parties in this litigation. Certain Defendants have abused ambiguities of this Court's order and produced disks that are unusable in the form in which they were provided. (See Affidavit of Roy Platell, attached and incorporated as exhibit "A." Mr. Plattel's affidavit details the problems with Defendants' production). Plaintiff has provided disks

that are immediately usable without any editing or programming required to make them useful. This is not true of certain Defendants' production.

2. INDUSTRY STANDARD DATA EXCHANGE CONVENTIONS

Plaintiff requests that the following, industry standard conventions be required of the document databases produced by all parties to this action, in order that all computerized document indices produced in this action are equally useful to all opposing counsel. The industry standard format for delivery of such data is:

ASCII, comma delimited text, with:

1. field names identified
2. field types identified
3. field lengths identified
4. dates in MM/DD/YY format throughout the entire database
5. identification of the fields' order
6. consistent field order throughout the entire database

Data supplied in other forms present serious problems, and is usually either unusable, or useable only after extensive hand conversion that may or may not yield a database that contains the same information as the originating database ordered to be produced.

3. PROBLEMS CREATED BY PRODUCTION OF NON-STANDARD DATA AND EXAMPLES OF EACH

Examples of some of the following problems encountered with certain Defendants' production of computerized document indices include the following:

No field names. Without field names Plaintiff is left with cryptic tables of information such as "030886" and is asked to figure out if this stands for March 8th, 1986, and whether March 8th is the date the document was created, or reviewed, or added to the database. Perhaps this is not a date at all, but a document ID number. A person's name may be either the recipient or author of

the document. Without field names Plaintiff cannot know for certain what this data is, and the information is useless.

Fields not comma delimited. The industry standard data format for the exchange of data between different systems is called “ASCII comma delimited”. In ASCII comma delimited format each field is separated by a comma, and each record is separated by a line feed. For example:

Data arriving in ASCII comma delimited form looks like this:

"001219","John Q. Smith","Memo on meeting","12/25/87","Albert Allan"¶

"001220","John Q. Smith","Letter to File","12/26/87","Legal Issues File"¶

A database program can import data in this form automatically, producing fields and records to look like this:

001219	John Q. Smith	Memo on meeting	12/25/87	Albert Allan
001220	John Q. Smith	Letter to File	12/26/87	Legal Issues File

The vertical columns of information are called “fields,” they represent a single category of data. In the above example, the fields are, “document number”, “author”, “subject” etc. from left to right. “Records” refers to the horizontal rows of data. In our example above each record is a separate document that is being indexed.

Fields and records can be sorted, searched and otherwise made to reveal useful information. The database program relies on the commas, quotation marks and linefeeds to divide the data into fields and records. If the data arrives with fields delimited by a character other than a comma these characters must be changed to commas. Likewise, if a quotation mark or carriage return is missing or misplaced the data will not import into a database correctly. These missing or misplaced characters must be found and replaced individually by hand, an expensive and time

consuming process. (For an example see exhibit “C”).

No field type. This refers to the type of information contained in the field. Database programs require the user to specify what “type” of information is in each field. Examples of field types include “text”, “date”, “number”, “time”, “calculation”. This information is necessary when importing data into a database. Plaintiff must import the data supplied by Defendants into a database in order for it to be in a usable form. Due to the volume of the information it is useless unless it can be sorted and searched in a database.

Dates in wrong format. The standard format recognized by most database programs is the standard US system of displaying month, date, and year in the form MM/DD/YY. In this form January 31st, 1997 would read “01/31/97”. In order for a database to recognize this number as a date it must be in this order, and remain in this format consistently throughout the entire database. Dates in other forms are not recognized and must be converted by hand to the standard format before the data can be imported to a usable form (in a database). (For an example see exhibits “B” and “D”).

Other problems with Defendants’ production are far more obvious failures to comply with this Court’s orders. Many of the disks produced are completely blank, or are otherwise completely useless. Examples of these problems include:

No data on disk. A blank disk with a label stating that a file full of data is contained on the disk is similar to the production of blank pieces of paper.

Completely jumbled data. Disks arriving in this condition are of no use to Plaintiff, as there is no way that these disks can be viewed or used by anyone. They are clearly not in compliance with this courts orders, and it is unfair to allow production of material in this manner by some parties while other parties produce data that is immediately useful. Plaintiff files this

motion pursuant to Fed. R. Civ. P. 1, “to secure the just, speedy, and inexpensive determination” of this case. The document databases have been determined by this Court to be important tools necessary for the efficient preparation of this case for trial, and all parties are entitled to equally useful forms of this data. (For an example see exhibit “E”).

4. PRODUCTION ON DISKETTES

Plaintiff further moves the Court for an Order modifying its prior requirement that these indices be produced on diskettes. Due to the large size of these databases it would be less costly and time consuming for all parties if the Court ordered producing parties to confer on the data storage medium prior to production. If no agreement is reached on the data storage medium the producing party should be required to produce data on 3.5 inch, 1.44 megabyte IBM format diskettes, for all production of five or fewer diskettes. For databases that require more than five diskettes, production should be on CD's, in CD-R, ISO 9660 format, if there is no agreement on an alternate data storage medium between the parties prior to the date production is due.

5. PAST PRODUCTION

For disks and databases previously produced by parties that do not comply with the requirements of this motion Plaintiff suggests that the party be given ten (10) days to produce databases in a form consistent with this motion.

6. CONCLUSION

For the foregoing reasons, Plaintiff, the State of Texas, respectfully requests the Court to compel Defendants to disclose document databases in the industry standard format described above.

Respectfully submitted:

DAN MORALES

ATTORNEY GENERAL OF TEXAS
Texas Bar No.: 14417450

JORGE VEGA

First Assistant Attorney General
Texas Bar No.: 20533800

JAVIER AGUILAR

Special Assistant Attorney General
Texas Bar No.: 00936300

TOM PERKINS

Chief, Consumer Protection Division
Texas Bar No.: 15790850

HARRY G. POTTER III

Special Assistant Attorney General
Texas Bar No.: 16175300
P.O. Box 12548
Capitol Station
Austin, Texas 78711-2548
512.463.2191
512.463.2063 Fax

JOHN M. O'QUINN, P.C.

440 Louisiana St., Suite 2300
Houston, Texas 77002
713.223.1000
713.222.6903 Fax
Texas Bar No.: 15296000

JOHN EDDIE WILLIAMS JR.

8441 Gulf Freeway, Suite 600
Houston, Texas 77017
713.649.6464
713.943.6226 Fax
Texas Bar No.: 21600300

WAYNE REAUD
REAUD, MORGAN & QUINN, INC.
801Laurel
Beaumont, Texas 77701
409.838.1000
409.833.8236 Fax
Texas Bar No.: 16642500

HAROLD W. NIX
NIX LAW FIRM
205 Linda Drive
P. O. Box 679
Daingerfield, Texas 75638
903.645.7333
903.645.5389 Fax
Texas Bar No.: 15041000

HUGH E. MCNEELY, of Counsel for
Provost & Umphrey Law Firm, L.L.P.
2901 Turtle Creek Drive, Suite 201
Port Arthur, Texas 77642
409.727.0800
409.727.7671 Fax
Louisiana Bar No.: 10,628

GRANT KAISER
KAISER & MORRISON, P.C.
Suite 1440 Lyric Centre
440 Louisiana Street
Houston, Texas 77002-1634
713.223.0000
713.223.440 Fax
State Bar No.: 11078900

WALTER UMPHREY, P.C.*
490 Park
P. O. Box 4905
Beaumont, Texas 77074
409.835.6000
409.838.8888 Fax
Texas Bar No.: 20380000
ATTORNEY-IN-CHARGE

* By Grant Kaiser, with permission.

CERTIFICATE OF SERVICE

I hereby certify compliance with Fed. R. Civ. P. 5 and Case Management Order of November 5, 1996, that a true a correct copy of the foregoing document and diskette has been sent by overnight delivery service and filed on February 25, 1997, to the following:

ADMINISTRATIVE LIAISON COUNSEL FOR ALL DEFENDANTS:

Howard Waldrop
Atchley, Russell, Waldrop & Hlavinka, L.L.P.
1710 Moores Lane
P. O. Box 5517
Texarkana, TX 75505-5517
903.792.8246
903.792.5801 Fax

Respectfully submitted,

Grant Kaiser

**IN THE UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF TEXAS
TEXARKANA DIVISION**

THE STATE OF TEXAS,	§	CIVIL NO.:	5:96-CV-0091
Plaintiff,	§		
	§		
VS.	§		
	§	JUDGE: DAVID FOLSOM	
THE AMERICAN TOBACCO	§		
COMPANY; R. J. REYNOLDS	§		
TOBACCO COMPANY;	§	MAGISTRATE JUDGE:	
BROWN & WILLIAMSON TOBACCO	§	WENDELL C. RADFORD	
CORPORATION; B. A. T.	§		
INDUSTRIES, P. L. C.; PHILIP	§		
MORRIS, INC.;	§		
LIGGETT GROUP, INC.; LORILLARD	§		
TOBACCO COMPANY, INC.; UNITED	§		
STATES TOBACCO COMPANY; HILL	§		
& KNOWLTON, INC.; THE	§		
COUNCIL FOR TOBACCO.	§		
RESEARCH-USA, INC. (Successor to	§		
Tobacco Institute Research Committee);	§		
and THE TOBACCO INSTITUTE, INC.	§		
Defendants.	§		

ORDER COMPELLING DISCLOSURE

Plaintiff's Motion To Compel Document Databases In A Usable Form, filed on February 25, 1996, was considered by the Court today. After considering all filings related to this motion, arguments, if any, of counsel and applicable law, the Court is of the opinion the Motion should be granted.

IT IS THEREFORE ORDERED that Defendants, except B.A.T. Industries, P.L.C., produce to Plaintiff all databases in the following form:

ASCII, comma delimited text, with:

1. field names identified
2. field types identified
3. field lengths identified

4. dates in MMDDYY format throughout the entire database
5. identification of the fields' order
6. consistent field order throughout the entire database

IT IS FURTHER ORDERED that parties producing computerized data confer with the opposing administrative liaison counsel on the data storage medium to be used prior to their production. If no agreement is reached on the data storage medium the producing party shall produce the data on 3.5 inch, 1.44 megabyte IBM format diskettes, for all production of five or fewer diskettes. For databases that require more than five diskettes, production should be on CD's, in CD-R, ISO 9660 format, if there is no agreement on an alternate data storage medium between the parties prior to the date production is due.

SIGNED THIS _____ DAY OF _____, 1997.

MAGISTRATE JUDGE WENDELL C. RADFORD
UNITED STATES DISTRICT COURT