

**UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
OAKLAND DIVISION**

JASON TRABAKOOLAS and SHEILA
STETSON, individually and on behalf of all
others similarly situated,

Plaintiffs,

v.

WATTS WATER TECHNOLOGIES, INC.,
WATTS REGULATOR CO., WATTS
ANDERSON-BARROWS METAL CORP.,
WATTS PLUMBING TECHNOLOGIES
(TAIZHO) CO., LTD., SAVARD PLUMBING
COMPANY, WOLVERINE BRASS, INC.,
AND JOHN DOES 1-100,

Defendants.

Case No. 4:12-cv-01172-YGR

CLASS ACTION

**~~[PROPOSED]~~ STIPULATED
PROTOCOL FOR PRODUCTION OF
DOCUMENTS AND
ELECTRONICALLY STORED
INFORMATION**

Honorable Yvonne Gonzalez Rogers

The parties to this action agree to the following protocol for the production of documents and electronically stored information (“ESI”):

I. SCOPE AND APPLICATION

This Protocol supplements the provisions of Federal Rules of Civil Procedure 16, 26, 33, 34, 37 and 45 insofar as those Rules relate to the production of documents and ESI. To the extent that instructions in Requests for Production conflict with this Protocol, the Protocol shall govern.

II. DEFINITIONS

A. **“Documents”** shall be interpreted broadly and shall include all writings, drawings, graphs, charts, photographs, and other data, data records or data compilations – stored in any medium from which information can be obtained. The parties shall negotiate in good

faith on the scope of documents that will be deemed responsive to each side's discovery requests.

B. **“Electronically stored information” or “ESI,”** as used herein, means and refers to computer generated information or data of any kind, stored in or on any storage media located on computers, file servers, disks, tape or other real or virtualized devices or media. Non limiting examples of ESI include:

1. Digital communications (e.g., e-mail, voice mail, instant messaging);
2. Contents of E-Mail server stores (e.g., Lotus Domino .NSF or Microsoft Exchange .EDB)
3. Applicable contents of databases (e.g., Access, Oracle, SQL Server data, SAP);
4. Word processed documents (e.g., Word or WordPerfect files and drafts);
5. Spreadsheets and tables (e.g., Excel or Lotus 123 worksheets);
6. Accounting application data (e.g., QuickBooks, Money, Peachtree data);
7. Image and Facsimile files (e.g., .PDF, .TIFF, .JPG, .GIF images);
8. Video and Animation (e.g., .AVI and .MOV files);
9. Presentations (e.g., PowerPoint, Corel Presentations)
10. Computer Aided Design/Drawing Files; and
11. Cloud based or other virtualized ESI, including application, infrastructure and data where applicable

C. **“Cloud computing”** means user accessing a remote computer hosted by a cloud service provider over the Internet or intranet to access software programs to create, save, or retrieve data, for example, to send messages or create documents, spreadsheets, or databases.

D. **“Native data format”** means the format of ESI in which it was normally created, viewed and/or modified by the producing party in the usual course of its business and in its regularly conducted activities.

E. **“Metadata”** means information about information, or data about data, and

includes (i) information embedded in or associated with a native file that is not ordinarily viewable or printable from the application that generated, edited, or modified such native file which describes the characteristics, origins, usage and/or validity of the electronic file, and/or (ii) information generated automatically by the operation of a computer or other information technology system when a native file is created, modified, transmitted, deleted or otherwise manipulated by a user of such system.

F. **“Static image”** means a representation of ESI produced by converting a Native File into a standard image format capable of being viewed and printed on standard computer systems. A Tagged Image File Format (TIFF) image is an example of a Static Image.

G. **“Forensic image”** (mirror image) of a hard drive or other storage device. A process that preserves the entire contents of a hard drive or other storage device by creating a bit-by-bit copy of the original data without altering the original media.

H. **“Extracted text”** The text of a native file extracted during ESI processing of the native file, most commonly when native files are converted to TIFF format.

I. **“Document unitization”** is the process of determining where a document begins (its first page) and ends (its last page), with the goal of accurately describing what was a “unit” as it was received by the party or was kept in the ordinary course of business by the document’s custodian. A “unit” includes attachments, for example, an email with an attached spreadsheet. Physical unitization utilizes actual objects such as staples, paper clips and folders to determine pages that belong together as documents. Logical unitization is the process of human review of each individual page in an image collection using logical cues to determine pages that belong together as documents.

J. **“Parent - child relationships”** Related documents are described as having a parent/child relationship, for example, where the email is the parent and an attached spreadsheet is the child.

K. **“OCR” (Optical Character Recognition)** is a process that converts a picture of text into searchable text. The quality of the created text can vary greatly depending on the quality of the original document, the quality of the scanned image, the accuracy of the recognition software and the quality control process of the provider.

L. **“MD5 Hash”** is an algorithm that is used to verify data integrity through the creation of a 128-bit message digest from data input that is as unique to that specific data as a fingerprint is to a specific individual.

M. **“PDF”** (Portable Document Format) is a file format created by Adobe that allows a range of options, including electronic transmission, viewing, and searching.

N. **“Media”** means an object or device, real or virtualized, including but not limited to a disc, tape, computer or other device, on which data is or was stored.

III. PROCESSING OF COLLECTED DATA

A. **Time Zone.** When processing ESI, **Eastern Standard Time** should be selected as the time zone. To the extent that a party has already processed ESI using a different time zone, the producing party will note the time zone used in its processing. Additionally, the use of daylight savings time will be employed.

B. **Deduplication.** Parties will perform global de-duplication on identical ESI as follows:

1. **Electronic Files:** Electronic files will be de-duplicated based upon calculated MD5 Hash values for binary file content. File contents only will be used for MD5 Hash value calculation and will not include operating system metadata (filename, file dates) values.

2. **Messaging Files:** Messaging files will be de-duplicated based upon MD5 Hash values for the message family, including parent object and attachments. The following fields will be used to create the unique value for each message: To; From; CC; BCC; Date Sent; Subject; Body; and, MD5 Hash values for all attachments, in attachment order. Messaging materials will be de-duplicated at a family level, including message and attachment(s).

3. For all files processed, a DupCustodian field will be populated with the duped out custodians if known and the DupCustodian field should be included at time of final production.

C. **Date Range.** Parties will agree upon a date range to apply to the processing of the data that will eliminate data outside of the range from further review and production.

D. **Compressed files.** Compression file types (i.e., .CAB, .GZ, .TAR, .Z, .ZIP) shall be decompressed in a reiterative manner to ensure that a zip within a zip is decompressed into the lowest possible compression resulting in individual folders and/or files.

IV. SEARCH METHODOLOGY

A. The Parties will cooperate in good faith regarding the disclosure and formulation of appropriate search terms and protocols in advance of any search to collect ESI. With the objective of limiting the scope of review and production, and thereby reducing discovery burdens, the parties shall meet and confer as early as possible, and in advance of commencing

any production efforts, to develop a list of agreeable search terms and disclose semantic synonyms. Semantic synonyms shall mean without limitation code words, terms, phrases or illustrations, acronyms, abbreviations, or non-language alphanumeric associational references to relevant ESI, or information that may lead to relevant ESI.

B. Once a final search protocol has been agreed to and executed, a requesting Party may, in good faith, seek to expand the scope of the search. Where such a request is made, the Parties will meet and confer and attempt in good faith to reach agreement as to the timing and conditions of such expansion. If the Parties cannot reach agreement, the dispute may be presented to the Court pursuant to the Court's Standing Order in Civil Cases regarding discovery disputes.

C. The parties will continue to meet and confer regarding any search process issues as necessary and appropriate. This ESI protocol does not address or resolve any other objections to the scope of the parties' respective discovery requests.

V. FORMAT OF PRODUCTION

A. **Document Image Format.** Unless otherwise required by the terms of this ESI protocol (Section V.(B)) or agreed to in writing by a requesting party, ESI is to be produced in TIFF format. Accompanying this TIFF shall be a multipage text (.TXT) file containing searchable text from the native file, and the metadata as discussed later in this document. The parties are not obligated to populate manually any of the fields in Appendix A if such fields cannot be extracted from a document, with the exception of the CUSTODIAN, and VOLUME, which shall be populated by the producing party. Load files of the static images should be created and produced together with their associated static images to facilitate the use of the produced images by a document management or litigation support database system. The parties shall meet and confer to the extent reasonably necessary to facilitate the import and use of the produced materials with commercially available document management or litigation support

software. Each party agrees to bear their own costs for the conversion of native files to TIFF and the production, storage, and maintenance of ESI in TIFF format. Nothing in this ESI Protocol shall prevent a party from seeking relief from the Court pursuant to Rule 26(b)(2)(B) and/or Rule 26(c).

B. Native Data Format. Unless redacted or otherwise agreed to in writing by a requesting party, Microsoft Excel, Microsoft Powerpoint, Microsoft Access Database and .txt file extensions shall be produced in native data format - as the ESI exists on the producing party's computer system.

C. Production of Structured Data. To the extent a response to discovery requires production of discoverable electronic information contained in a database, and where data is reasonably accessible, the parties shall meet and confer to agree upon a set of queries to be made for discoverable information.

D. Production of Physical Documents. Documents or records which either were originally generated or created as ESI but now only exist in physical hard-copy format, or documents or records that were originally generated in hard-copy format shall be converted to a single page TIFF file and produced following the same protocols set forth herein or otherwise agreed to by the parties.

E. Document Unitization. For files not produced in their native data format, each page of a document shall be electronically saved as an image file. If a document consists of more than one page, the unitization of the document and any attachments and/or affixed notes shall be maintained as it existed in the original when creating the image files.

The producing party shall produce a unitization file ("load file") for all produced Documents in accordance with the following formatting, including a separate load file for native Documents:

1. **OCR and Extracted Text Files (.TXT Files):**
 - a. Single text file per document containing all the document's pages.
 - b. Pages separated by form feed character (decimal 12, hex 0xC)
 - c. Filenames should be of the form:
<Bates num>.txt
Where <Bates num> is the BATES number of the first page in the document.
 - d. Text must be encoded in UTF-8.
2. **Images Files:**
 - a. Group IV TIFF at 300 X 300 dpi resolution.
 - b. 8 1/2 X 11 inch page size, except for documents requiring different resolution or page size.
 - c. Single page per image.
 - d. Single image per file.
 - e. TIFF is default FORMAT unless the following formats are agreed to: jpeg, jpeg2000, giff, png, single image tiff, and bmp.
 - f. Filenames should be of the form:
<Bates num>.<ext>
Where <Bates num> is the BATES number of the page, and <ext> is the appropriate extension for the image format (.jpg, .tif, .png, etc)
3. **Index Files:**
 - a. Comma Separated Value (.CSV) files (commonly .DAT files).

- b. First line must contain the column/field names (set forth in Paragraph 1(c) herein).
- c. Every row must have the same number of columns/fields (empty values are acceptable).
- d. Text must be encoded in UTF-8.
- e. Values must be enclosed by decimal 17, hex 0x11 (ascii device control 1).
- f. Values must be separated by decimal 19, hex 0x13 (ascii device control 3).

4. **Native Index Files:**

- a. .CSV files.
- b. All lines contain data – first row must NOT contain column headers.
- c. Every row must have 2 columns/fields (empty values are NOT acceptable).
- d. First column/field must contain the BATES number for the document.
- e. Second column/field must contain the filename (NOT the full path) of the native file. Filenames must be unique in the production – unless the content is identical (for example, a native and an OCR text file may both be named XYZ01234567.TXT if they are identical).
- f. Text must be encoded in UTF-8.
- g. Values must be enclosed by double quotes (ascii character 34).
- h. Values must be separated by a comma (ascii character 44).

F. **Exception Files:** The parties will make reasonable efforts to resolve any user-generated files that have been collected that cannot be produced and/or imaged.

G. **Color.** For files not produced in their native format, if an original document contains color, the producing party may produce black and white image(s). At the request of the receiving party, the parties shall meet and confer to discuss production of color image(s) for specific documents.

H. **Bates Numbering and Other Unique Identifiers.** For files not produced in their native data format, each page of a produced document shall have a legible, unique page identifier (“Bates Number”) electronically “burned” onto the TIFF image in such a manner that information from the source document is not obliterated, concealed, or interfered with. There shall be no other legend or stamp placed on the document image unless the document has been redacted in accordance with applicable law, court order or pursuant to an applicable protective order or other case specific principle. In the case of redacted materials, a designation may be “burned” onto the document’s image at a location that does not obliterate or obscure any information from the source document. Additionally, any natively produced files must have a Bates number added to the file name: i.e. if a file is named “example.pdf” then when producing it would be “PREFIX00000001_example.pdf. Bates numbers must meet the criteria below:

1. be unique across the entire document production,
2. maintain a constant length (zero/0-padded) across the entire production,
3. contain no special characters or embedded spaces, and
4. be sequential within a given document.
5. If a Bates number or set of Bates numbers is skipped in a production, and not otherwise identified on a privilege log, the producing party will disclose the Bates numbers or ranges in a cover letter accompanying the production.

I. **Production Media.** Documents shall be produced on CD-ROM, DVD, external hard drive (with standard PC compatible interface), or such other readily accessible computer or electronic media as the parties may hereafter agree upon (the “Production Media”). Each item of Production Media shall include: (1) text referencing that it was produced in *Trabakoolas, et al. v. Watts Water Technologies, Inc., et al.*, (Case No. 12-cv-01172-YGR), (2) the production date, and (3) the Bates number range of the materials contained on such Production Media item. Load

files shall not span across media (e.g., CDs, DVDs, Hard Drives, etc.); a separate volume shall be created for each piece of media delivered.

J. **Cover Letter.** A cover letter shall accompany each transmission of ESI discovery providing basic information including the number of media, the unique identifiers of the media, a brief description of the contents, any applicable Bates ranges or other unique production identifiers, and any necessary passwords to access the content. Passwords should not be in the cover letter accompanying the data, but in a separate communication.

K. **Electronic Text Files.** For files not produced in their native data format, text files for produced documents shall be produced reflecting the full text that has been electronically extracted from the original, native electronic files (“Extracted Text”). The Extracted Text shall be provided in ASCII text format. The text files will be named with the unique Bates number of the first page of the corresponding document followed by the extension “.txt.” Electronic text must be extracted directly from the native electronic file unless the document was redacted, an image file, or a hard copy document. In these instances, a text file shall be created using OCR and shall be produced in lieu of extracted text. Additionally, a list of Bates numbers of files without text will be provided upon the completion of the final production.

L. **Hash Value.** The parties will either preserve the MD5 hash value on the electronic documents or preserve the MD5 hash value and produce this hash value with the electronic document.

M. **Metadata.** The Metadata fields identified in Appendix A will be extracted and produced (to the extent available and not privileged) in an attached file accompanying the TIFF productions. This list of fields does not create any obligation to create or manually code fields that are not automatically generated by the processing of the ESI, that do not exist as part of the original Metadata of the document, or that would be burdensome or costly to obtain. The parties

retain the right to move the Court for the production of additional electronic metadata fields should ongoing discovery reveal the need for such Metadata.

N. **Attachments.** Email attachments and embedded files or links, to the extent produced, should be mapped to their parent by the Document or Production number. If attachments and embedded files are combined with their parent documents, then “BeginAttach” and “EndAttach” fields listing the unique beginning and end number for each attachment or embedded document must be included.

O. **Embedded Objects.** Embedded objects will not be automatically extracted. The parties will confer about specific embedded materials that will be produced and if they are unable to agree they can bring this matter to the Court.

P. **Compressed Files.** Compression file types (i.e., .CAB, .GZ, .TAR, .Z, .ZIP) shall be decompressed in a reiterative manner to ensure that a zip within a zip is decompressed into the lowest possible compression resulting in individual folders and/or files. (See III- F, “File Inclusion List.”)

Q. **Dynamic Fields.** Documents with dynamic fields for file names, dates, and times will be processed to show the field code (e.g., “[FILENAME]” or “[AUTODATE]”), rather than the values for such fields existing at the time the file is processed.

VI. OBJECTIONS TO ESI PRODUCTION

A. Documents that present imaging or format production problems shall be promptly identified and disclosed to the requesting party; the parties shall then meet and confer to attempt to resolve the problems.

B. If either party objects to producing the requested information on the grounds that such information is not reasonably accessible because of undue burden or cost, or because production in the requested format is asserted to be not reasonably accessible because of undue

burden or cost, the responding party will inform the requesting party of the format in which it is willing and able to produce, the nature and location of the information claimed to not be reasonably accessible, the reason(s) why the requested form of production would impose an undue burden or is unreasonably costly, and afford the requesting party 10 business days from receipt of such notice to propose an alternative means of compliance with the request.

C. If a party believes that responsive ESI no longer exists or is not retrievable, the responding party shall identify the ESI and explain the circumstances surrounding the change in its status and the reasons the ESI no longer exists or is no longer retrievable. The requesting party can schedule a meet and confer to discuss whether any backup or copy of the responsive ESI exists, and if so, whether and how best to obtain responsive information.

VII. PRIVILEGE AND WORK PRODUCT CLAIMS

For any document withheld in its entirety or produced but redacted, the producing party will produce privilege/redaction logs in Excel format or any other format that permits electronic sorting and searching. The Parties shall have no obligation to log privileged communications with their clients concerning this litigation or privileged information generated after the commencement of this litigation. A privilege log will be produced 45 days after completion of the production of the first set of custodial files and 45 days after any further set of custodial files, or no less than five days prior to a particular custodian's deposition for which custodial documents were withheld, whichever is sooner. In no event, however, will any privilege log need to be produced prior to December 15, 2012.

When there is a chain of privileged e-mails, the producing party need only include one entry on the privilege/redaction log for the entire e-mail chain and need not log each e-mail contained in the chain separately. The producing party and the receiving party may modify the deadlines for production of privilege/redaction logs by agreement.

For each document withheld or redacted, the producing party's privilege/redaction logs shall include the following information:

- (a) Custodian or source,
- (b) Date,
- (c) Author(s),
- (d) For documents produced, but redacted on the ground of privilege, the starting and ending Bates number;
- (e) Recipient(s), CC(s) and BCC(s) (for e-mail and hard-copy communication such as letters and internal memoranda);
- (f) Specification of the privilege claimed;
- (g) A description of the document and the basis for the privilege claim.

Following the receipt of a privilege/redaction log, a receiving party may identify, in writing, specific documents that it believes require further explanation. If a party challenges a request for further information, the parties shall meet and confer to try to reach a mutually agreeable solution prior to informing the Court of unresolved issues.

Nothing herein is intended to modify or supersede any Protective Orders entered in this matter.

VIII. CATEGORIES OF ESI DISCOVERABLE ONLY UPON A SHOWING OF GOOD CAUSE

The following categories of ESI are discoverable only upon a showing of good cause: (1) “deleted,” “slack,” “fragmented,” or “unallocated” data on hard drives; (2) random access memory (RAM) or other ephemeral data; (3) on-line access data such as temporary internet files, history, cache, cookies, etc.; (4) data in metadata fields that are frequently updated automatically, such as last-opened dates; (5) backup and archival data that is substantially duplicative of data that is more accessible elsewhere; (6) the entire E-Mail Server Stores; (7) sound recordings (such as WAV and .MP3 files); (8) entire databases; (9) contact and relationship management data; (10) calendar or diary application data; (11) online access data (e.g., Temporary Internet Files,

History, Cookies); (12) network access and server activity logs; and (13) project management application data.

IX. CONTINUING OBLIGATIONS OF ALL PARTIES TO MEET AND CONFER

To expedite discovery of relevant electronic evidence and reduce each parties' costs, the parties' attorneys and IT professionals will informally cooperate and discuss procedures or protocols to facilitate identification, retrieval and production of computerized information. The parties will work with one another in good faith to resolve any issues, disputes or objections that arise in connection with electronic discovery issues before raising such matters with the court. Issues shall be raised promptly in writing, and the parties shall have good faith discussions to attempt to resolve the matter. The responsibility shall be continuing, unless otherwise ordered by the Court.

The parties shall comply with the Court's Standing Order regarding discovery disputes, including the in-person meet and confer requirement.

DAVID S. MacCUIISH
TODD BENOFF
LINDSAY G. CARLSON
ALSTON & BIRD LLP

DATED: October __, 2012

DAVID S. MacCUIISH
Attorneys for Defendant Watts Regulator Co. and
Specially Appearing Defendant Watts Water
Technologies, Inc.

JOSEPH J. TABACCO, JR.
TODD A. SEAVER
BERMAN DEVALERIO

STEVE W. BERMAN
ANTHONY D. SHAPIRO
JEFF D. FRIEDMAN
ELAINE T. BYSZEWSKI

HAGENS BERMAN SOBOL SHAPIRO LLP

DONALD L. PERELMAN
GERARD A. DENVER
FINE, KAPLAN AND BLACK, P.C.

SIMON BAHNE PARIS
PATRICK HOWARD
CHARLES J. KOCHER
SALTZ, MONGELUZZI, BARRETT &
BENDESKY, P.C.

DATED: October __, 2012

SIMON PARIS
Attorneys for Plaintiffs and the proposed Classes

**Pursuant to Stipulation (Dkt. No. 95) and subject to the Court's modifications
herein, IT IS SO ORDERED. This Order terminates Dkt. No. 95.**

DATED: October 22, 2012



**Honorable Yvonne Gonzalez Rogers
United States District Court Judge**

APPENDIX A

TABLE OF METADATA FIELDS¹

<u>Field Name</u>	<u>Format</u>	<u>Description</u>
BegDoc	Text	Starting Bates Number for record
EndDoc	Text	Ending Bates Number for record
Pages	Numeric	Number of pages in document
BegAttach	Text	Bates number, first page of parent document for the document group
EndAttach	Text	Bates number, last page of last attachment, for the document group
Custodian	Text	Custodian
Volume	Text	Volume of Production
DateCreate	Date, MM/DD/YYYY	Creation date of original file
FileExt	Text	Original document file extension
Subject	Text	Subject line of email
Author	Text	Document author
From	Text	Author of message
To	Text	Recipients of message

¹ These metadata fields will be provided to the extent it is available in the native file at the time of collection. Some fields may have limited utility. For example, where a template is used to generate a document, the author on the face of the document would differ from that in the metadata. Similarly, the file extension may not be indicative of the file type, as users can modify the extension without affecting the way the file is read by the application.

CC	Text	CC recipients of message
BCC	Text	BCC recipients of message
DateSent	Date, MM/DD/YYYY	Sent date if email message
Designation	Text	Confidentiality designation
MD5_Hash	Hash	Digital footprint of the file
Redacted	Y or N	“Yes” or “No” indication of whether the Document at issue is redacted.
DupCustodian	Text	List of duplicate custodians for each file; Please note, custodians listed more than once in any given email file, such as in the “From” and “CC” fields will have the custodian list in both the regular “Custodian” field as well as the “DupCustodian” field.