

# UELMA Preservation

---

Jason Judt & Daniel Kruse

Office of the Revisor of Statutes, Minnesota

# UELMA System Overview

- Server-side authentication
  - Complete versioning system
  - Tracking and auditing tools developed
  - Support for multiple formats (only PDF is authentic)
  - Long term preservation strategy
-

# Authentication

- What gets authenticated
    - Documents containing legal material: Chapters, Sections
    - ~~Publication front matter, tables, indexes (book aids)~~
  - Every published PDF is fingerprinted (secure hash)
  - Upload PDFs to web site to verify authenticity
  - Pros:
    - Easy
    - Cheap
    - No third party dependencies
  - Cons:
    - Requires manual action
    - Requires internet connection
-

# THE OFFICE OF THE REVISOR OF STATUTES

## Document Authentication

Authenticated PDF copies of legislative documents are considered official electronic records under the [Uniform Electronic Legal Material Act](#).

All material viewed on the Office of the Revisor, State of Minnesota website is sent via secure servers (SSL) with encryption. If you need an official record or require additional verification of authenticity, follow the steps below.

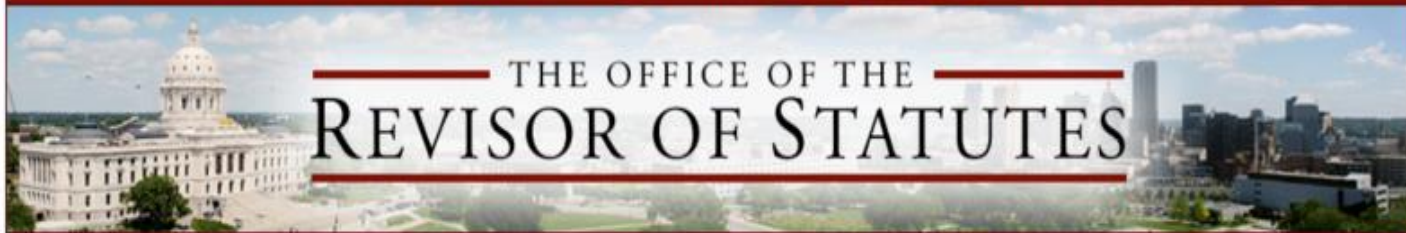
- Step 1. Find the PDF for the law, statute, or administrative rule you wish to authenticate: [Rule Part 6264.0400](#)
- Step 2. Download that PDF into the folder on your own computer in which you store files
- Step 3. Click on the "Browse/Choose File" button below, find the downloaded file, and click on it
- Step 4. Click on the "Authenticate" button below

Choose File No file chosen

Authenticate Return

# Authentication

Upload Page



## Document Authentication

[Authenticate](#)

**Your file "3E.03.pdf" is an authentic PDF copy of document**

Minnesota [Statute Section 3E.03](#) as published 9/02/2015

[Details](#)

Authenticated PDF copies of legislative documents are considered official electronic records under the [Uniform Electronic Legal Material Act](#).

All material viewed on the Office of the Revisor, State of Minnesota website is sent via secure servers ([SSL](#)) with encryption. If you need an official record or require additional verification of authenticity, follow the steps below.

Step 1. Find the PDF for the law, statute, or administrative rule you wish to authenticate: [Rule Part 6264.0400](#)

Step 2. Download that PDF into the folder on your own computer in which you store files

Step 3. Click on the "Browse/Choose File" button below, find the downloaded file, and click on it

Step 4. Click on the "Authenticate" button below

Choose File No file chosen

[Authenticate](#) [Return](#)

# Authentication

Success



## Document Authentication

### Your file "output.pdf" is NOT authenticated

Some of the reasons that your PDF document could not be authenticated include:

- The PDF document may have been altered \*
- The uploaded file is not a PDF document
- This document did not originate from the Office of the Revisor [website](#)
- It is not an official version of the Minnesota Constitution, Laws, Statutes or Rule covered by the [Uniform Electronic Legal Materials Act](#)

\* **Note to users of the Firefox web browser:** older versions of Firefox have been known to modify PDF documents, which makes the authentication process fail. We recommend either using a different web browser or updating your Firefox browser, then re-downloading the PDF to obtain an authentic copy.

Authenticated PDF copies of legislative documents are considered official electronic records under the [Uniform Electronic Legal Material Act](#).

All material viewed on the Office of the Revisor, State of Minnesota website is sent via secure servers (SSL) with encryption. If you need an official record or require additional verification of authenticity, follow the steps below.

Step 1. Find the PDF for the law, statute, or administrative rule you wish to authenticate: [Rule Part 6264.0400](#)

Step 2. Download that PDF into the folder on your own computer in which you store files

Step 3. Click on the "Browse/Choose File" button below, find the downloaded file, and click on it

Step 4. Click on the "Authenticate" button below

Choose File No file chosen

Authenticate Return

# Authentication

## Failure

# Versioning – before UELMA

- MN Administrative Rules are amended continuously
  - Changes to rules are immediately published to the web (static PDF/HTML foldered by year)
  - A full MN Rule publication is published every odd year
  - A single rule supplement book is published every even year
  - Versions could get lost (!!!)
-

# Versioning

- Continuous timeline maintained for each published document
  - Versions created with respect to publication dates – not effective dates
  - Extra version metadata: Publish Reason (ex. Agency Action, Legislative Action, Editorial Update)
  - Erroneous publications are tracked and treated the same as other publications, but they are hidden from the version list
-



# THE OFFICE OF THE REVISOR OF STATUTES

Retrieve by number Rules Statutes Laws Rules Court Rules Constitution Revisor's Office Search Law by Keyword

Rules > Natural Resources Department > Chapter 6264 > Part 6264.0400

6264.0300

## Minnesota Administrative Rules

PDF Authenticate

### 6264.0400 DESIGNATED SPECIAL MANAGEMENT WATERS.

Subpart 1. [Repealed, 28 SR 629]

Subp. 2. **Catch and release fishing on trout streams.** While on or fishing in the following waters, angling for trout shall be limited to catch and release only, whereby any trout caught must be immediately returned to the water. Catch and release angling for trout shall be legal during the open season for trout in inland streams. Legal methods for taking trout are limited to artificial lures and flies.

A.	Name	Location	County
	Kedron Creek	T104, R. 13, S.36 from the confluence with Bear Creek upstream approximately 1.1 miles to the permanent source - entire stream	Fillmore
	Root River, South Fork	T102, R. 8, S. 17, 18, 19 and R. 9, S. 24, 25, 26, from the County Road 12 crossing in T102, R. 8, S. 17, upstream approximately 7.7 miles to the permanent source (large springs) in T102, R. 9, S. 26	Fillmore
	Whitewater River, Middle Branch	T107, R. 10, S. 19, 20, 30 and R. 11, S. 24, 25, 26, 35 and T106, Olmsted and R. 11, S. 2, 3, 10, from the posted boundary at the point where Winona a state park trail crosses the stream in the SE 1/4 NE 1/4 S. 19, T107, R. 10, upstream approximately 9.2 miles to the source in the SW 1/4 NE 1/4, S. 10, T106, R. 11	

Subp. 3. [Repealed, 39 SR 5]

### Subp. 4. **Mille Lacs Lake and associated tributaries special management regulations.**

A. While a person is on or fishing in Mille Lacs Lake or its associated tributaries to the posted boundaries, all northern pike in possession must be less than 24 inches in length or greater than 36 inches in length. All northern pike that are 24 to 36 inches in length, inclusive, must be immediately returned to the water. A person's possession limit may not include more than one northern pike over 36 inches in length.

B. The possession limit for smallmouth bass is one. While a person is on or fishing in Mille Lacs Lake or its associated tributaries to the posted boundaries, all smallmouth bass in possession must be 21 inches

#### Resources

About Minnesota Rules

Search Minnesota Rules

Agency List

Rules Index (topics)

Minnesota Rule Status

State Register

Invalid Rules

Chapter 6264

Table of Parts

Full Chapter Text

Part 6264.0400

Version List

# Versioning

Rule Text Page

Example: Minnesota Rule 6264.0400

# THE OFFICE OF THE REVISOR OF STATUTES

6264.0300

## Minnesota Administrative Rules

[View All Versions](#) [Pdf](#) [Authenticate](#)

### 6264.0400 Versions

Date Published Electronically	PDF	Details
<a href="#">3/03/2016</a>	<a href="#">Pdf</a>	Agency Action
<a href="#">1/26/2016</a>	<a href="#">Pdf</a>	Agency Action
<a href="#">8/28/2015</a>	<a href="#">Pdf</a>	Legislative Action
<a href="#">3/02/2015</a>	<a href="#">Pdf</a>	Agency Action
<a href="#">1/28/2015</a>	<a href="#">Pdf</a>	Agency Action
<a href="#">7/14/2014</a>	<a href="#">Pdf</a>	Agency Action
<a href="#">3/05/2014</a>	<a href="#">Pdf</a>	Agency Action
<a href="#">12/09/2013</a>	<a href="#">Pdf</a>	Electronically Published in compliance with <a href="#">MN Statutes 3E</a>

Copyright © 2016 by the Revisor of Statutes, State of Minnesota. All rights reserved.

**Resources**

- [About Minnesota Rules](#)
- [Search Minnesota Rules](#)
- [Agency List](#)
- [Rules Index \(topics\)](#)
- [Minnesota Rule Status](#)
- [State Register](#)
- [Invalid Rules](#)
- [Chapter 6264](#)
- [Table of Parts](#)
- [Full Chapter Text](#)
- [Part 6264.0400](#)
- [Version List](#)

# Versioning

Rule Version List

Example: Minnesota Rule 6264.0400

MN Legislature Get Connected About the Legislature Mobile Site Search the MN Legislature

House Senate Joint Schedules Legislators Committees Bills Law Multimedia Publications

THE OFFICE OF THE REVISOR OF STATUTES

Retrieve by number Statutes Statutes Laws Rules Court Rules Constitution Revisor's Office Search Law by Keyword

## Document Authentication

[Authenticate](#)

### You have authenticated a historical version

Minnesota [Rule Part 6264.0400](#) as published 1/28/2015

[Details](#)

See [Current version](#) of Minnesota Rule Part 6264.0400

Authenticated PDF copies of legislative documents are considered official electronic records under the [Uniform Electronic Legal Material Act](#).

All material viewed on the Office of the Revisor, State of Minnesota website is sent via secure servers ([SSL](#)) with encryption. If you need an official record or require additional verification of authenticity, follow the steps below.

Step 1. Find the PDF for the law, statute, or administrative rule you wish to authenticate: [Rule Part 6264.0400](#)

Step 2. Download that PDF into the folder on your own computer in which you store files

Step 3. Click on the "Browse/Choose File" button below, find the downloaded file, and click on it

Step 4. Click on the "Authenticate" button below

Choose File No file chosen

Authenticate Return

# Versioning

Authenticating older version of rule

Example: Minnesota Rule 6264.0400

# Tracking and auditing

- Challenge: UELMA documents are permanent
    - Mistakes live forever
    - Interested in reducing the amount of republishing
  - Solution #1: Backend tracking of publishing
    - Every step for every document is audited as it moves through publishing code
    - Results are stored for 6 months and then consolidated/summarized
  - Solution #2: Constructed an intermediate preview area (web sandbox)
    - Staff can review a published document for correctness
    - If satisfied staff can “approve” the publication, making the publication permanent
-

# Formats

- Store and fingerprint PDF and XML
    - PDF is authenticatable
    - XML is not made available and is not authenticatable at this time
  - HTML is transformed on-the-fly from XML
    - Allows documents to be displayed with modern techniques
    - Allows future features to be integrated into HTML easily
-

# Preservation

- Approach #1: KEEPS
    - WORM-based duplicate store of documents
    - Regular validation that live documents match the archive
  - Approach #2: Dark archive at the Minnesota Historical Society
    - Documents placed in the vault, never touched again
-

# KEEPS Goals

- Preserve
  - Legal documents must be preserved in perpetuity.
- Validate
  - Documents available to the public are authentic.
- Authenticate
  - Users can check the authenticity of documents they possess.

*Minnesota Statutes*

*3E.07*

*Subdivision 2. **Requirements if preservation in electronic form.***

*If legal material is preserved under subdivision 1 in an electronic record, the official publisher shall:*

*(1) ensure the integrity of the record;*

*(2) provide for backup and disaster recovery of the record; and*

*(3) ensure the continuing usability of the material.*

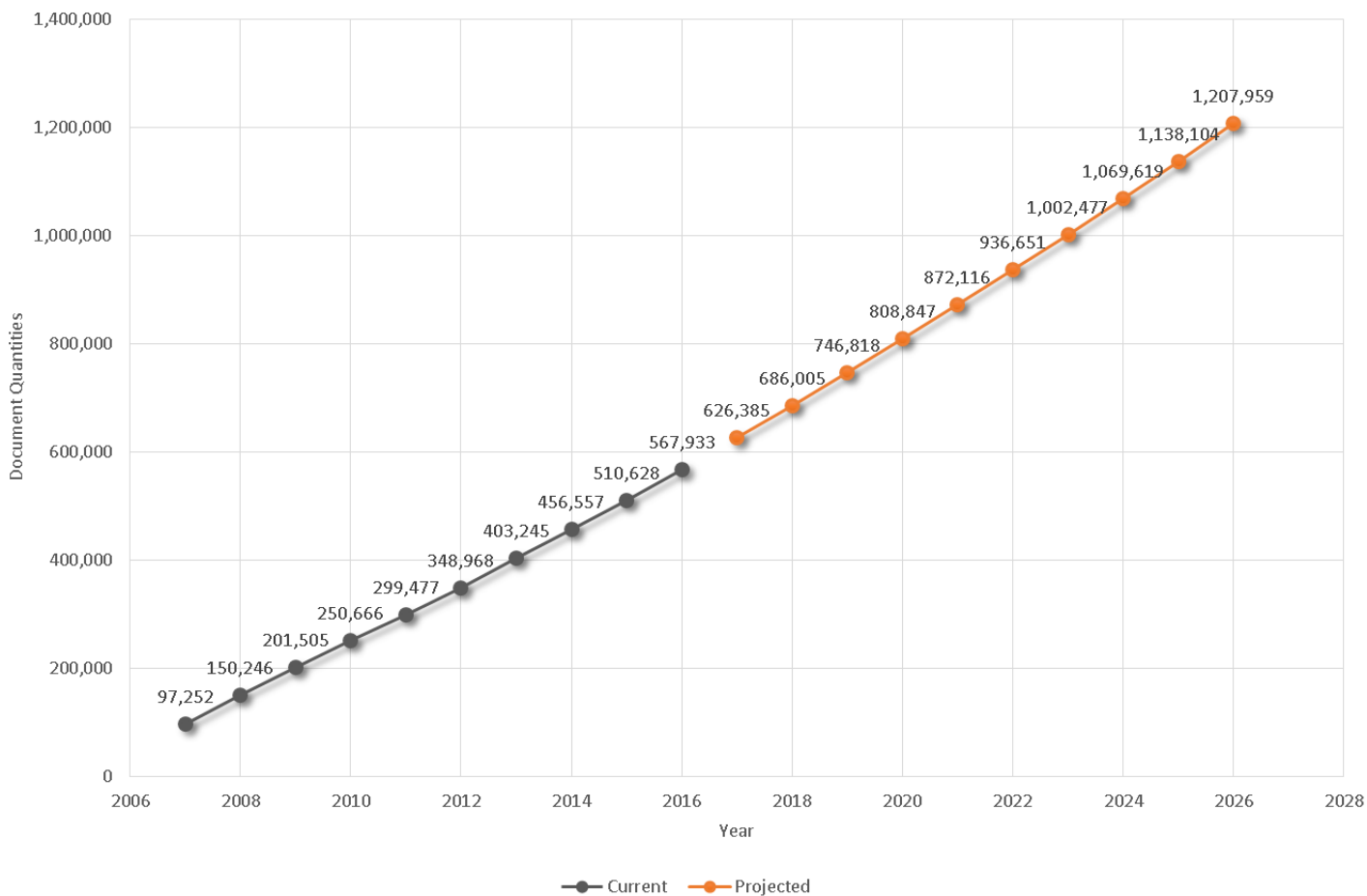
---

# Requirements

- Speed
    - Validate all UELMA documents daily.
  - Independence
    - Must run autonomously without affecting other legislative information systems.
  - Accuracy
    - Must detect any change to a document down to bit level manipulation.
  - Monitoring
    - Email notifications of validation errors.
  - Sleuthing
    - A web interface combining all available information.
-



### Document Estimates



# Archive Projections

510k UELMA Documents

54k Additions in 2016 increasing at 2% annually

We project the size of our repository will double in the next 8 years

# Alpha Prototype

- Requirements Gathered.
  - Cloud based, utilizing Amazon Web Services
  - Robust and scalable
-

# Beta Prototype

- WORM (write once read many) disks identified as an ideal solution.
  - Cloud services don't allow user owned hardware.
  - Developed on existing in house servers and processing.
  - Proof of concept established.
  - Time estimates for nightly validation determined via scaled testing.
-

# Release Candidate

- Refined requirements.
  - Determined primary technologies.
    - Oracle 12c
      - Current PDF's are stored in Oracle database
      - Native SHA2 hashing
      - Minimize data IO
    - KEEPS server
      - GreenTec WORM disks
        - Hardware enforced write once.
      - Windows Server 2012 R2
      - Java se8
-

## KEEPS Server

### Schedule Manager

Validator



Archiver



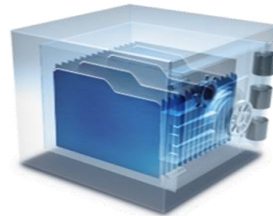
WORM Contents Populator



Scheduled tasks: Must not run concurrently.  
Queue-able if another task is running.



User initiated tasks



KEEPS Document Repository  
(WORM Disk)

BulkArchiver



# Server Architecture

# Process Flow – Archival and Validation

## Archive

- A newly published document is detected
- The document is archived to the worm drive

## Validation

- The worm drive is analyzed and compared to the database.
  - Discrepancies are noted in an error list.
  - Failures are reported via Email and the Intranet
-

# Testing

- Individual Corruption tests
    - Unauthorized document inserted into database
    - A document removed from the database
    - Changes to an existing document in the database
  - Load tests
    - Archival of 50k Statute pdf's (6.3GB) occurs in < 37 minutes.
    - Validation of 500k pdf's (65GB) occurs in < 30 minutes.
-

# Summary

- On Site
  - We control our data.
  - Low implementation cost
  - No reliance on outside services.
  - Deployed to production at the beginning of September.
-