Bryce pointed out that we have many incorrect checksumAlgorithm strings various MNs.  See [https://github.nceas.ucsb.edu/KNB/arctic-data/issues/283](https://github.nceas.ucsb.edu/KNB/arctic-data/issues/283). The upshot is that SHA-* is the broadly supported syntax.

I checked the strings with:

```java
package org.dataone.tests;
import java.security.MessageDigest;
import java.security.NoSuchAlgorithmException;
import java.util.ArrayList;
import java.util.List;
public class MessageDigestDTest {
    public static void main(String[] args) {
        MessageDigest md = null;
        List<String> algorithms = new ArrayList<String>();
algorithms.add("MD5");
algorithms.add("MD-5");
algorithms.add("SHA1");
algorithms.add("SHA-1");
algorithms.add("SHA224");
algorithms.add("SHA-224");
algorithms.add("SHA256");
algorithms.add("SHA-256");
algorithms.add("SHA384");
algorithms.add("SHA-384");
algorithms.add("SHA512");
algorithms.add("SHA-512");
        for (String algorithm : algorithms) {
            try {
                md = MessageDigest.getInstance(algorithm);
                System.out.println(md.getAlgorithm() + " is recognized.");
            } catch (NoSuchAlgorithmException e) {
                System.out.println(e.getMessage());
            }
        }
    }
}
```

and got:

MD5 is recognized.
MD-5 MessageDigest not available
SHA1 is recognized.
SHA-1 is recognized.
SHA224 MessageDigest not available
SHA-224 is recognized.
SHA256 MessageDigest not available
SHA-256 is recognized.
SHA384 MessageDigest not available
SHA-384 is recognized.
SHA512 MessageDigest not available
SHA-512 is recognized.

Change MNodeService, CNodeService, and D1NodeService methods that send or receive SystemMetadata documents and validate the given string with MessageDigest.getInstance(algorithm). If we get a NoSuchAlgorithmException exception, throw an InvalidSystemMetadata exception for the call.

History
#1 - 12/20/2017 09:31 AM - Matt Jones

The definition of the [ChecksumAlgorithm](https://releases.dataone.org/online/api-documentation-v2.0.1/apis/Types.html#Types.ChecksumAlgorithm) type says that algorithm names must be drawn from the Library of Congress controlled vocabulary:

> The cryptographic hash algorithm used to calculate a checksum. DataONE recognizes the Library of Congress list of cryptographic hash algorithms that can be used as names in this field, and specifically uses the madsrdf:authorityLabel field as the name of the algorithm in this field. See: Library of Congress Cryptographic Algorithm Vocabulary. All compliant implementations must support at least SHA-1 and MD5, but may support other algorithms as well.

We should be checking against that list, and not the Java names, which may not be language neutral.

#2 - 01/17/2018 04:10 PM - Jing Tao

According the list here [http://id.loc.gov/vocabulary/preservation/cryptographicHashFunctions.html](http://id.loc.gov/vocabulary/preservation/cryptographicHashFunctions.html) some names from the list are:

- MD5
- SHA-1
- SHA-256
- SHA-384
- SHA-512

It doesn't show SHA-224. I am not sure if it is in the list.