Q: Why are you messing with Solr?
A: DSpace 6 uses Solr 4, which was first released in 2012, went on limited support in early 2015, and has been EOL since about mid-2016 with the release of Solr 6. It seemed unwise not to stay up-to-date with recent improvements and bug-fixes in Solr.

Q: So, did anything change?
A: Solr is now a prerequisite rather than a dependency. You will need to install and start it separately.

Q: Why did DSpace change from a dependency (built with DSpace) to a prerequisite?
A: Solr’s packaging was reorganized. Starting with Solr 5, Solr can only be built as a stand-alone application with its own embedded Servlet container. It could not be deployed in the same way as Solr 4, so changes were needed anyway.

Q: Didn’t DSpace make modifications to Solr? How do we do that now?
A: DSpace made two tiny modifications: (a) a filter to allow connections only from the local host, and (b) a listener to force the use of DSpace’s logging configuration at startup. These modifications have been dropped from the DSpace 7 kit.

Stock Solr has its own logging configuration which is easy to set up.

The code for both of these modifications can still be found in Github if you want it.

Q: Is it hard to set up Solr myself? Will you give step-by-step instructions?
A: Installing Solr can be as simple as unpacking a Zip archive. You will need to arrange to have it started when DSpace is running. If you have local standards for application layout, you can probably adapt Solr to them.

No, we will not provide step-by-step instructions, because different OSes have different steps, and because the Solr maintainers know the product better. There is extensive documentation at http://lucene.apache.org/solr/resources.html

At least one site has contributed to the DSpace Wiki some notes on how they deploy Solr, specific to their OS and local requirements.

Q: We already have DSpace with Solr 4. How do we get to Solr 7?
A: DSpace still ships with empty Solr cores, now configured for Solr 7. You will need to copy these to the place where Solr expects to find them. Instructions for doing that are in the DSpace upgrade documentation.

The ‘search’ and ‘oai’ cores must be re-indexed, using tools provided with DSpace.

The ‘authority’ and ‘statistics’ cores contain information not available elsewhere. They must be dumped from Solr 4 and restored to Solr 7, using tools provided with DSpace.

Q: Why so complicated?
A: There have been a lot of changes between Solr 4 and Solr 7! The Solr field analyzers which DSpace has used are no longer part of Solr 7. The indexes must be rebuilt using the current set of analyzers.

Q: We have sharded our ‘statistics’ core. Does that make a difference?
A: TBD

Q: Are there other changes I should know about?
A: A stock Solr deployment uses port 8983, not 8080. You will need to ensure that your DSpace configuration has URLs appropriate to the new Solr setup. It is possible that you will need to adjust firewall rules to permit contact between DSpace and Solr. (This also gives you new opportunities to control access to Solr separately from access to DSpace.)

Q: So, are there any advantages for my site?
A: Perhaps. You can deploy Solr and DSpace on separate hosts, and tune them separately for best performance. You can have different firewall rules for them. If you already have a big Solr deployment, you can simply add DSpace’s cores to it and save some duplication.

You may consider it an advantage that you can now upgrade Solr and DSpace independently.