Ligatures and ZWJ

For most scripts, originally including Arabic script, the ZERO WIDTH JOINER (ZWJ) is intended to increase the connectedness of adjacent glyphs. The following summary has been in The Unicode Standard (TUS) essentially unchanged since Version 3.1, but is here quoted from Version 14.0, Section 23.2:

> In other words, given the three broad categories below, ZWJ requests that glyphs in the highest available category (for the given font) be used:
> 1. Ligated
> 2. Cursively connected
> 3. Unconnected[1]

However, due to a number of issues, this rule does not apply to Arabic script. All versions of TUS since 4.0 have included this exception to the above rule:

> For backward compatibility, between Arabic characters a ZWJ acts just like the sequence <ZWJ, ZWNJ, ZWJ>, preventing a ligature from forming instead of requesting the use of a ligature that would not normally be used. **As a result, there is no plain text mechanism for requesting the use of a ligature in Arabic text.**[2]
> [emphasis added]

Therefore it is inappropriate for a font developer to attempt to devise plain-text mechanisms for authors to request optional ligature formation.

For those who want to do some research, you might look at:

- http://www.unicode.org/L2/L2000/00023.txt for the reasons why ligature encouragement was added to the function of ZWJ
- http://www.unicode.org/L2/L2002/02363-nelson-zwj-zwnj.pdf for reasons the exception for Arabic script was added in TUS 14.0

---

1 The Unicode Consortium. The Unicode Standard, Version 14.0.0, (Mountain View, CA: The Unicode Consortium, 2021. ISBN 978-1-936213-29-0), pg 906.

2 ibid., pg 908.