1. Сразу скажу – правильнее рассматривать как единую единицу, а не два разночастеречных (наречие+глагол) слова. Это вводное словосочетание. Вообще, в речи каждого конкретного человека вводными словами и дискурсивными словами могут быть разные единицы, поэтому надо определять это отдельно для каждого говорящего. Смысл вышесказанного такой, что выделить отдельный маркер для вводных и дискурсивных слов было бы разумно. Пока в дз отметила как наречие и глагол.
2. Как раз – если «как» еще можно определить как местоимение, то найти отдельный маркер для «раз» как-то очень сложно. Это тоже стоит объединить для разметки в один токен – DET. В домашней разметке я оба слова токенизировала как DET.
3. Это – может быть DET может быть PRON. В некоторых случаях люди вставляют «это» просто так, с каким-то дискурсивным значением и не совсем по правилам грамматики. Но обычно это происходит для подчеркивания какой-то фразы, так что в дз я отмечала как DET.
4. Это все – должно рассматриваться как одна единица, отмечала оба слова как DET.
5. Поэтому – даже в корпусе разметка зависит от контекста, кажется, что в корпусе представлены не все возможные варианты – есть местоимение и союз, также возможен вариант наречие.