

Expected Results

Data Set Version: 1-1, Revised 09-27-19

Tested Tool: _____

Search Engine: _____

You can print this document and use it to record the results of each test case to paper as an intermediate step to recording the results in the String Search test suite.

MS Windows Data Set				
Case/String	Encoding	Active	Deleted	Unalloc
FT-SS-01				
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
FT-SS-02				
Wolf	ascii	0864 0866 0868	0865 0867 0869	0870
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
WOLF	ascii	0880 0882 0884	0881 0883 0885	0886
WereWolf	ascii	0912 0914 0916	0913 0915 0917	0918
wolf	ascii	0848 0850 0852	0849 0851 0853	0854
FT-SS-03				
Wolf	ascii	0864 0866 0868	0865 0867 0869	0870
WOLF	ascii	0880 0882 0884	0881 0883 0885	0886
wolf	ascii	0848 0850 0852	0849 0851 0853	0854
FT-SS-04				
Panda and fox	ascii	2944 2946 2948	2945 2947 2949	
FT-SS-05				
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
WereWolf	ascii	0912 0914 0916	0913 0915 0917	0918
FT-SS-06				
fox and not tiger				
	ascii	0784 0792 0800	0788 0796 0804	
	utf-16-be	0787 0795 0803	0791 0799 0807	
	utf-16-le	0786 0794 0802	0790 0798 0806	
	utf-8	0785 0793 0801	0789 0797 0805	
FT-SS-07-CJK-char				
中国				
	utf-16-be	1987 1995 2003	1991 1999 2007	2011
	utf-16-le	1986 1994 2002	1990 1998 2006	2010
	utf-8	1985 1993 2001	1989 1997 2005	2009
東京				
	utf-16-be	2883 2891 2899	2887 2895 2903	2907
	utf-16-le	2882 2890 2898	2886 2894 2902	2906
	utf-8	2881 2889 2897	2885 2893 2901	2905
FT-SS-07-CJK-hangul				
서울				
	utf-16-be	2371 2379 2387	2375 2383 2391	2395

Windows Data Set

MS Windows Data Set				
Case/String	Encoding	Active	Deleted	Unalloc
	utf-16-le	2370 2378 2386	2374 2382 2390	2394
	utf-8	2369 2377 2385	2373 2381 2389	2393
FT-SS-07-CJK-kana				
みつびし				
	utf-16-be	2307 2315 2323	2311 2319 2327	2331
	utf-16-le	2306 2314 2322	2310 2318 2326	2330
	utf-8	2305 2313 2321	2309 2317 2325	2329
スバル				
	utf-16-be	2179 2187 2195	2183 2191 2199	2203
	utf-16-le	2178 2186 2194	2182 2190 2198	2202
	utf-8	2177 2185 2193	2181 2189 2197	2201
FT-SS-07-Cyrillic				
Сибирь				
	utf-16-be	1187 1195 1203	1191 1199 1207	1211
	utf-16-le	1186 1194 1202	1190 1198 1206	1210
	utf-8	1185 1193 1201	1189 1197 1205	1209
FT-SS-07-Latin				
Schönheit				
	utf-16-be	2499 2507 2515	2503 2511 2519	2523
	utf-16-le	2498 2506 2514	2502 2510 2518	2522
	utf-8	2497 2505 2513	2501 2509 2517	2521
garçon				
	utf-16-be	2691 2699 2707	2695 2703 2711	2715
	utf-16-le	2690 2698 2706	2694 2702 2710	2714
	utf-8	2689 2697 2705	2693 2701 2709	2713
FT-SS-07-NoBOM				
QuarterHorse				
	ascii	2560 2568 2576	2564 2572 2580	2584
	utf-16-be	2563 2571 2579	2567 2575 2583	2587
	utf-16-le	2562 2570 2578	2566 2574 2582	2586
	utf-8	2561 2569 2577	2565 2573 2581	2585
Россия				
	utf-16-be	1059 1067 1075	1063 1071 1079	1083
	utf-16-le	1058 1066 1074	1062 1070 1078	1082
	utf-8	1057 1065 1073	1061 1069 1077	1081
فلافل				
	utf-16-be	1539 1547 1555	1543 1551 1559	1563
	utf-16-le	1538 1546 1554	1542 1550 1558	1562
	utf-8	1537 1545 1553	1541 1549 1557	1561
中國				
	utf-16-be	1923 1931 1939	1927 1935 1943	1947
	utf-16-le	1922 1930 1938	1926 1934 1942	1946
	utf-8	1921 1929 1937	1925 1933 1941	1945
FT-SS-07-Norm				
Mäuse (NFD)				
	utf-16-be	3283 3291 3299	3287 3295 3303	3307
	utf-16-le	3282 3290 3298	3286 3294 3302	3306
	utf-8	3281 3289 3297	3285 3293 3301	3305
Mäuse (NFC)				
	utf-16-be	3091 3099 3107	3095 3103 3111	3115
	utf-16-le	3090 3098 3106	3094 3102 3110	3114
	utf-8	3089 3097 3105	3093 3101 3109	3113

Windows Data Set

MS Windows Data Set				
Case/String	Encoding	Active	Deleted	Unalloc
infinity (No Ligature)				
	ascii	3344 3352 3360	3348 3356 3364	3368
	utf-16-be	3347 3355 3363	3351 3359 3367	3371
	utf-16-le	3346 3354 3362	3350 3358 3366	3370
	utf-8	3345 3353 3361	3349 3357 3365	3369
infinity (Ligature)				
	utf-16-be	3411 3419 3427	3415 3423 3431	3435
	utf-16-le	3410 3418 3426	3414 3422 3430	3434
	utf-8	3409 3417 3425	3413 3421 3429	3433
libertà (NFD)				
	utf-16-be	3219 3227 3235	3223 3231 3239	3243
	utf-16-le	3218 3226 3234	3222 3230 3238	3242
	utf-8	3217 3225 3233	3221 3229 3237	3241
libertà (NFC)				
	utf-16-be	3027 3035 3043	3031 3039 3047	3051
	utf-16-le	3026 3034 3042	3030 3038 3046	3050
	utf-8	3025 3033 3041	3029 3037 3045	3049
mañana (NFD)				
	utf-16-be	3155 3163 3171	3159 3167 3175	3179
	utf-16-le	3154 3162 3170	3158 3166 3174	3178
	utf-8	3153 3161 3169	3157 3165 3173	3177
mañana (NFC)				
	utf-16-be	2963 2971 2979	2967 2975 2983	2987
	utf-16-le	2962 2970 2978	2966 2974 2982	2986
	utf-8	2961 2969 2977	2965 2973 2981	2985
FT-SS-07-RTL				
الكسكس				
	utf-16-be	1475 1483 1491	1479 1487 1495	1499
	utf-16-le	1474 1482 1490	1478 1486 1494	1498
	utf-8	1473 1481 1489	1477 1485 1493	1497
FT-SS-08-Email				
berlin@ deutschland.net	ascii	1280 1282 1284	1281 1283 1285	1286
iron.man@ marvel.com				
	ascii	0944 0952 0960	0948 0956 0964	0968
	utf-16-be	0947 0955 0963	0951 0959 0967	0971
	utf-16-le	0946 0954 0962	0950 0958 0966	0970
	utf-8	0945 0953 0961	0949 0957 0965	0969
kgb@moscow. red.square.ru	ascii	1264 1266 1268	1265 1267 1269	1270
potus@capitol. gov	ascii	1248 1250 1252	1249 1251 1253	1254
FT-SS-08-Phone				
(901)555-1111	ascii	1296 1298 1300	1297 1299 1301	1302
202.555.3270	ascii	1392 1394 1396	1393 1395 1397	1398
301.555-9009				
	ascii	1328 1336 1344	1332 1340 1348	1352
	utf-16-be	1331 1339 1347	1335 1343 1351	1355
	utf-16-le	1330 1338 1346	1334 1342 1350	1354

Windows Data Set

MS Windows Data Set				
Case/String	Encoding	Active	Deleted	Unalloc
	utf-8	1329 1337 1345	1333 1341 1349	1353
800-555-1122	ascii	1312 1314 1316	1313 1315 1317	1318
FT-SS-08-SS				
123-45-6789	ascii	1008 1010 1012	1009 1011 1013	1014
987-65-4321	ascii	1024 1026 1028	1025 1027 1029	1030
999-55-1321	ascii	1040 1042 1044	1041 1043 1045	1046
FT-SS-09-Doc				
crossbow	utf-8	8007 9007		8515 9515
flintlock	utf-8	8001 9001		8509 9509
longbow	utf-8	8006 9006		8514 9514
nitroglycerin				
	UTF	9005		9513
	utf-8	8005		8513
peroxide				
	UTF	9004		9512
	utf-8	8004		8512
revolver				
	utf-16-be	9002		9510
	utf-8	8002		8510
rifle	utf-8	8000 9000		8508 9508
shotgun				
	utf-16-be	9003		9511
	utf-8	8003		8511
FT-SS-09-Frag				
California	utf-8	6000		
Washington	utf-8	6006		
FT-SS-09-Lost				
SecretKey				
	ascii			7000
	utf-16-be			7003
disconnected				
	ascii			7016
	utf-16-le			7018
FT-SS-09-MFT				
bear				
	ascii	7008	7012	
	utf-16-be	7011	7015	
	utf-16-le	7010	7014	
	utf-8	7009	7013	
FT-SS-09-Meta				
cañón	utf-8	2625 2633 2641	2629 2637 2645	2649
thunderbird	ascii	0928 0930 0932	0929 0931 0933	0934
FT-SS-09-Stem				
cities				
	ascii	0480 0488 0496	0484 0492 0500	0504
	utf-16-be	0483 0491 0499	0487 0495 0503	0507
	utf-16-le	0482 0490 0498	0486 0494 0502	0506
	utf-8	0481 0489 0497	0485 0493 0501	0505
city	ascii	0464 0466 0468	0465 0467 0469	0470
knife	ascii	0576 0578 0580	0577 0579 0581	0582
knives	ascii	0592 0594 0596	0593 0595 0597	0598
plan	ascii	0144 0146 0148	0145 0147 0149	0150

Windows Data Set

MS Windows Data Set				
Case/String	Encoding	Active	Deleted	Unalloc
plans	ascii	0160 0162 0164	0161 0163 0165	0166
planned	ascii	0176 0178 0180	0177 0179 0181	0182
planner	ascii	0192 0194 0196	0193 0195 0197	0198
planning	ascii	0208 0210 0212	0209 0211 0213	0214
stealer	ascii	0736 0738 0740	0737 0739 0741	0742
stealing	ascii	0672 0674 0676	0673 0675 0677	0678
steals	ascii	0688 0690 0692	0689 0691 0693	0694
steal	Ascii	0656 0658 0660	0657 0659 0661	0662
FT-SS-10-Hex				
panda	ascii	2944 2946 2948	2945 2947 2949	2950
FT-SS-10-Regex				
DireWolf	ascii	0896 0898 0900	0897 0899 0901	0902
WereWolf	ascii	0912 0914 0916	0913 0915 0917	0918

DONE

UNIX Data Set

UNIX Data Set			
Case/String	Enc	Active	Deleted
FT-SS-01			
DireWolf	ascii	0904 0906 0908 0910	0905 0907 0909 0911
FT-SS-02			
Wolf	ascii	0872 0874 0876 0878	0873 0875 0877 0879
DireWolf	ascii	0904 0906 0908 0910	0905 0907 0909 0911
WOLF	ascii	0888 0890 0892 0894	0889 0891 0893 0895
WereWolf	ascii	0920 0922 0924 0926	0921 0923 0925 0927
wolf	ascii	0856 0858 0860 0862	0857 0859 0861 0863
FT-SS-03			
Wolf	ascii	0872 0874 0876 0878	0873 0875 0877 0879
WOLF	ascii	0888 0890 0892 0894	0889 0891 0893 0895
wolf	ascii	0856 0858 0860 0862	0857 0859 0861 0863
FT-SS-04			
panda and fox	ascii	2952 2954 2956 2958	2953 2955 2957 2959
FT-SS-05			
DireWolf	ascii	0904 0906 0908 0910	0905 0907 0909 0911
WereWolf	ascii	0920 0922 0924 0926	0921 0923 0925 0927
FT-SS-06			
fox and not tiger			
	ascii	0816 0824 0832 0840	0820 0828 0836 0844
	U-16-be	0819 0827 0835 0843	0823 0831 0839 0847
	U-16-le	0818 0826 0834 0842	0822 0830 0838 0846
	utf-8	0817 0825 0833 0841	0821 0829 0837 0845
FT-SS-07-CJK-char			
中国			
	U-16-be	2019 2027 2035 2043	2023 2031 2039 2047
	U-16-le	2018 2026 2034 2042	2022 2030 2038 2046
	utf-8	2017 2025 2033 2041	2021 2029 2037 2045
東京			
	U-16-be	2915 2923 2931 2939	2919 2927 2935 2943
	U-16-le	2914 2922 2930 2938	2918 2926 2934 2942
	utf-8	2913 2921 2929 2937	2917 2925 2933 2941
FT-SS-07-CJK-hangul			
서울			
	U-16-be	2403 2411 2419 2427	2407 2415 2423 2431
	U-16-le	2402 2410 2418 2426	2406 2414 2422 2430
	utf-8	2401 2409 2417 2425	2405 2413 2421 2429
FT-SS-07-CJK-kana			
みつびし			
	U-16-be	2339 2347 2355 2363	2343 2351 2359 2367
	U-16-le	2338 2346 2354 2362	2342 2350 2358 2366
	utf-8	2337 2345 2353 2361	2341 2349 2357 2365
スバル			
	U-16-be	2211 2219 2227 2235	2215 2223 2231 2239
	U-16-le	2210 2218 2226 2234	2214 2222 2230 2238
	utf-8	2209 2217 2225 2233	2213 2221 2229 2237
FT-SS-07-Cyrillic			

UNIX Data Set

UNIX Data Set			
Case/String	Enc	Active	Deleted
Сибирь			
	U-16-be	1219 1227 1235 1243	1223 1231 1239 1247
	U-16-le	1218 1226 1234 1242	1222 1230 1238 1246
	utf-8	1217 1225 1233 1241	1221 1229 1237 1245
FT-SS-07-Latin			
Schönheit			
	U-16-be	2531 2539 2547 2555	2535 2543 2551 2559
	U-16-le	2530 2538 2546 2554	2534 2542 2550 2558
	utf-8	2529 2537 2545 2553	2533 2541 2549 2557
garçon			
	U-16-be	2723 2731 2739 2747	2727 2735 2743 2751
	U-16-le	2722 2730 2738 2746	2726 2734 2742 2750
	utf-8	2721 2729 2737 2745	2725 2733 2741 2749
FT-SS-07-NoBOM			
QuarterHorse			
	Ascii	2592 2600 2608 2616	2596 2604 2612 2620
	U-16-be	2595 2603 2611 2619	2599 2607 2615 2623
	U-16-le	2594 2602 2610 2618	2598 2606 2614 2622
	utf-8	2593 2601 2609 2617	2597 2605 2613 2621
Россия			
	U-16-be	1091 1099 1107 1115	1095 1103 1111 1119
	U-16-le	1090 1098 1106 1114	1094 1102 1110 1118
	utf-8	1089 1097 1105 1113	1093 1101 1109 1117
فلافل			
	U-16-be	1571 1579 1587 1595	1575 1583 1591 1599
	U-16-le	1570 1578 1586 1594	1574 1582 1590 1598
	utf-8	1569 1577 1585 1593	1573 1581 1589 1597
中國			
	U-16-be	1955 1963 1971 1979	1959 1967 1975 1983
	U-16-le	1954 1962 1970 1978	1958 1966 1974 1982
	utf-8	1953 1961 1969 1977	1957 1965 1973 1981
FT-SS-07-Norm			
Mäuse (NFD)			
	U-16-be	3315 3323 3331 3339	3319 3327 3335 3343
	U-16-le	3314 3322 3330 3338	3318 3326 3334 3342
	utf-8	3313 3321 3329 3337	3317 3325 3333 3341
Mäuse (NFC)			
	U-16-be	3123 3131 3139 3147	3127 3135 3143 3151
	U-16-le	3122 3130 3138 3146	3126 3134 3142 3150
	utf-8	3121 3129 3137 3145	3125 3133 3141 3149
Infinity (No Ligature)			
	ascii	3376 3384 3392 3400	3380 3388 3396 3404
	U-16-be	3379 3387 3395 3403	3383 3391 3399 3407
	U-16-le	3378 3386 3394 3402	3382 3390 3398 3406
	utf-8	3377 3385 3393 3401	3381 3389 3397 3405
infinity (Ligature)			
	U-16-be	3443 3451 3459 3467	3447 3455 3463 3471
	U-16-le	3442 3450 3458 3466	3446 3454 3462 3470
	utf-8	3441 3449 3457 3465	3445 3453 3461 3469

UNIX Data Set

UNIX Data Set			
Case/String	Enc	Active	Deleted
Libertà (NFD)			
	U-16-be	3251 3259 3267 3275	3255 3263 3271 3279
	U-16-le	3250 3258 3266 3274	3254 3262 3270 3278
	utf-8	3249 3257 3265 3273	3253 3261 3269 3277
Libertà (NFC)			
	U-16-be	3059 3067 3075 3083	3063 3071 3079 3087
	U-16-le	3058 3066 3074 3082	3062 3070 3078 3086
	utf-8	3057 3065 3073 3081	3061 3069 3077 3085
Mañana (NFD)			
	U-16-be	3187 3195 3203 3211	3191 3199 3207 3215
	U-16-le	3186 3194 3202 3210	3190 3198 3206 3214
	utf-8	3185 3193 3201 3209	3189 3197 3205 3213
Mañana (NFC)			
	U-16-be	2995 3003 3011 3019	2999 3007 3015 3023
	U-16-le	2994 3002 3010 3018	2998 3006 3014 3022
	utf-8	2993 3001 3009 3017	2997 3005 3013 3021
FT-SS-07-RTL			
الكسكس			
	U-16-be	1507 1515 1523 1531	1511 1519 1527 1535
	U-16-le	1506 1514 1522 1530	1510 1518 1526 1534
	utf-8	1505 1513 1521 1529	1509 1517 1525 1533
FT-SS-08-Email			
berlin@deutschland.net	Ascii	1288 1290 1292 1294	1289 1291 1293 1295
iron.man@marvel.com			
	Ascii	0976 0984 0992 1000	0980 0988 0996 1004
	U-16-be	0979 0987 0995 1003	0983 0991 0999 1007
	U-16-le	0978 0986 0994 1002	0982 0990 0998 1006
	utf-8	0977 0985 0993 1001	0981 0989 0997 1005
kgb@moscow.red.square.ru	Ascii	1272 1274 1276 1278	1273 1275 1277 1279
potus@capitol.gov	Ascii	1256 1258 1260 1262	1257 1259 1261 1263
FT-SS-08-Phone			
(901)555-1111	ascii	1304 1306 1308 1310	1305 1307 1309 1311
202.555.3270	ascii	1400 1402 1404 1406	1401 1403 1405 1407
301.555-9009			
	ascii	1360 1368 1376 1384	1364 1372 1380 1388
	U-16-be	1363 1371 1379 1387	1367 1375 1383 1391
	U-16-le	1362 1370 1378 1386	1366 1374 1382 1390
	utf-8	1361 1369 1377 1385	1365 1373 1381 1389
800-555-1122	ascii	1320 1322 1324 1326	1321 1323 1325 1327
FT-SS-08-SS			
123-45-6789	ascii	1016 1018 1020 1022	1017 1019 1021 1023
987-65-4321	ascii	1032 1034 1036 1038	1033 1035 1037 1039
999-55-1321	ascii	1048 1050 1052 1054	1049 1051 1053 1055
FT-SS-09-Doc			
crossbow	utf-8	8515 9515	
flintlock	utf-8	8509 9509	
longbow	utf-8	8514 9514	

UNIX Data Set

UNIX Data Set			
Case/String	Enc	Active	Deleted
nitroglycerin			
	UTF	9513	
	utf-8	8513	
peroxide			
	UTF	9512	
	utf-8	8512	
revolver			
	U-16-be	9510	
	utf-8	8510	
rifle	utf-8	8508 9508	
shotgun			
	U-16-be	9511	
	utf-8	8511	
FT-SS-09-Meta			
cañón	utf-8	2657 2665 2673 2681	2661 2669 2677 2685
thunderbird	ascii	0936 0938 0940 0942	0937 0939 0941 0943
FT-SS-09-Stem			
cities			
	ascii	0512 0520 0528 0536	0516 0524 0532 0540
	U-16-be	0515 0523 0531 0539	0519 0527 0535 0543
	U-16-le	0514 0522 0530 0538	0518 0526 0534 0542
	utf-8	0513 0521 0529 0537	0517 0525 0533 0541
city	ascii	0472 0474 0476 0478	0473 0475 0477 0479
knife	ascii	0584 0586 0588 0590	0585 0587 0589 0591
knives	ascii	0600 0602 0604 0606	0601 0603 0605 0607
plan	ascii	0152 0154 0156 0158	0153 0155 0157 0159
plans	ascii	0168 0170 0172 0174	0169 0171 0173 0175
planned	ascii	0184 0186 0188 0190	0185 0187 0189 0191
planner	ascii	0200 0202 0204 0206	0201 0203 0205 0207
planning	ascii	0216 0218 0220 0222	0217 0219 0221 0223
stealer	ascii	0744 0746 0748 0750	0745 0747 0749 0751
stealing	ascii	0680 0682 0684 0686	0681 0683 0685 0687
steals	ascii	0696 0698 0700 0702	0697 0699 0701 0703
steal	ascii	0664 0666 0668 0670	0665 0667 0669 0671
FT-SS-10-Hex			
panda	ascii	2952 2954 2956 2958	2953 2955 2957 2959
FT-SS-10-Regex			
DireWolf	ascii	0904 0906 0908 0910	0905 0907 0909 0911
WereWolf	ascii	0920 0922 0924 0926	0921 0923 0925 0927